



A simple, efficient, mass-conservative approach to solving Richards' equation (openRE, v1.0)

Andrew M. Ireson^{1,2}, Raymond J. Spiteri³, Martyn P. Clark⁴, and Simon A. Mathias⁵

¹Global Institute for Water Security, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

²School of Environment and Sustainability, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

³Department of Computer Science, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

⁴Department of Geography and Planning, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

⁵Department of Engineering, Durham University, Durham, UK

Correspondence: Andrew M. Ireson (andrew.ireson@usask.ca)

Received: 17 July 2022 – Discussion started: 2 August 2022

Revised: 31 December 2022 – Accepted: 4 January 2023 – Published: 27 January 2023

Abstract. A simple numerical solution procedure – namely the method of lines combined with an off-the-shelf ordinary differential equation (ODE) solver – was shown in previous work to provide efficient, mass-conservative solutions to the pressure-head form of Richards' equation. We implement such a solution in our model openRE. We developed a novel method to quantify the boundary fluxes that reduce water balance errors without negative impacts on model runtimes – the solver flux output method (SFOM). We compare this solution with alternatives, including the classic modified Picard iteration method and the Hydrus 1D model. We reproduce a set of benchmark solutions with all models. We find that Celia's solution has the best water balance, but it can incur significant truncation errors in the simulated boundary fluxes, depending on the time steps used. Our solution has comparable runtimes to Hydrus and better water balance performance (though both models have excellent water balance closure for all the problems we considered). Our solution can be implemented in an interpreted language, such as MATLAB or Python, making use of off-the-shelf ODE solvers. We evaluated alternative SciPy ODE solvers that are available in Python and make practical recommendations about the best way to implement them for Richards' equation. There are two advantages of our approach: (i) the code is concise, making it ideal for teaching purposes; and (ii) the method can be easily extended to represent alternative properties (e.g., novel ways to parameterize the $K(\psi)$ relationship) and processes (e.g., it is straightforward to couple heat or solute transport), making it ideal for testing alternative hypotheses.

1 Introduction

Richards' equation (RE) describes the movement of water in variably saturated porous media. Almost any practical application of RE requires a numerical solution; yet RE remains challenging to solve reliably and accurately for a given set of boundary conditions and soil hydraulic properties (Farthing and Ogden, 2017). RE has been extensively reviewed (e.g., Vereecken et al., 2016; Farthing and Ogden, 2017, and references therein). RE has practical limitations in representing the flow processes in real soils containing macropores, especially with modeling infiltration and rapid percolation processes (Beven and Germann, 2013). The strength of RE is its ability to represent matrix drainage and capillary flows, which control evapotranspiration processes, and its ability to be coupled to heat and solute transport models. For this reason, RE remains a common approach to simulate soil moisture in many terrestrial system models (e.g., vadose zone models, ecohydrology models, and land-surface models; Vereecken et al., 2016; Clark et al., 2015, 2021).

Our objective in this paper is to (i) implement a simple and practical approach to solve RE that is efficient, is mass conservative, and uses open-source software (written in Python) that is readily available, including as a teaching tool; and (ii) present an improved mass balance calculation procedure for use with ordinary differential equation (ODE) solvers that apply adaptive time-stepping (ATS) schemes. We investigate how to maximize the efficiency and accuracy of ODE solvers and provide guidance on the subtle challenges that arise in

evaluating the boundary fluxes and the water balance. For our purposes, we define the following five criteria for success for an RE solver: (i) the solver should successfully reproduce benchmark solutions for $\psi(t, z)$ or $\theta(t, z)$; (ii) the solver should be mass conservative, with errors that are negligible based on the specific application; (iii) truncation errors in the simulated boundary fluxes should be negligible for the given time step; (iv) the solver should be computationally efficient; and (v) all else being equal on criteria (i)–(iv), the simplest code should be preferred. Simple code should be the most human readable/editable code, which means it should be clean, concise, modular, and free from redundancies.

The remainder of this paper is organized as follows. In Sect. 2, we describe a simple yet powerful approach to solve RE numerically using ODE solvers that we implement and test in the Python and MATLAB programming languages. In Sect. 2, we also discuss the complexities of closing the water balance in RE and present a novel solution that can be applied with any ODE solver: openRE. In Sect. 3, we benchmark the performance of our proposed solution against existing numerical models and solutions, including Hydrus 1D, and the solution method proposed by Celia et al. (1990). In Sect. 4, we summarize our recommendations.

2 Solving Richards' equation

RE is derived from the mass continuity equation applied to a control volume of soil, $\Delta x \Delta y \Delta z$ (L^3), and for one-dimensional vertical flow passing through the area $\Delta x \Delta y$ (L^2), we have

$$\frac{\partial m}{\partial t} = -\frac{\partial(q\rho)}{\partial z}, \quad (1)$$

where m (ML^{-3}) is the mass of water per control volume of soil, ρ (ML^{-3}) is the density of water, q (LT^{-1}) is the flux of water, and t (T) is time and z (L) is depth below some fixed datum. Assuming that density is constant, we may write

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z}, \quad (2)$$

where θ ($L^3 L^{-3}$) is the volumetric water content, defined as the volume of water per control volume of soil. The vertical flux is given by Darcy's law,

$$q_z = -K(\psi) \frac{dh}{dz} = -K(\psi) \left(\frac{d\psi}{dz} - 1 \right), \quad (3)$$

where h (L) and ψ (L) are the hydraulic head and matric potential head, respectively, and $K(\psi)$ (LT^{-1}) is the hydraulic conductivity. Combining Eqs. (2) and (3) we have

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left(K(\psi) \left(\frac{\partial \psi}{\partial z} - 1 \right) \right), \quad (4)$$

which is the mixed form of RE (Celia et al., 1990). If we let $C(\psi) = d\theta/d\psi$ (L^{-1}), we can write

$$C(\psi) \frac{\partial \psi}{\partial t} = \frac{\partial}{\partial z} \left(K(\psi) \left(\frac{\partial \psi}{\partial z} - 1 \right) \right), \quad (5)$$

which is the ψ -based form of RE (Celia et al., 1990). We can also express this as the θ form of RE, given by

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left(\frac{K(\theta)}{C(\theta)} \frac{\partial \theta}{\partial z} - K(\theta) \right), \quad (6)$$

where the constitutive relationships $C(\theta)$ and $K(\theta)$ are expressed as functions of θ (Celia et al., 1990). Changes in storage in Eqs. (4)–(6) are associated with the filling and draining of soil pores. Because $d\theta/d\psi = 0$ in saturated, or close to saturated, soils, an elastic storage term is needed to solve RE in these conditions. This term represents the compression of the pore water and the expansion of the pore space as a function of increasing pore water pressure (though the latter is orders of magnitude larger than the former). The ψ form of RE with elastic storage is written

$$\left(S_S \frac{\theta(\psi)}{\theta_s} + \frac{\partial \theta}{\partial \psi} \right) \frac{\partial \psi}{\partial t} = \frac{\partial}{\partial z} \left(K(\psi) \left(\frac{\partial \psi}{\partial z} - 1 \right) \right), \quad (7)$$

where S_S (L^{-1}) is the specific storage coefficient and θ_s is the saturated water content, equal to the porosity (Kavetski et al., 2001). If S_S is large, it may have a non-negligible impact on the water balance, which needs to be accounted for (e.g., as in Clark et al., 2021, Eq. 80). S_S is often treated as a numerical smoothing factor for RE when conditions are saturated or close to saturation, and its impact on the water balance is neglected (e.g., as in Tocci et al., 1997; Ireson et al., 2009). For convenience, hereafter we define $C(\psi) = S_S \frac{\theta(\psi)}{\theta_s} + \frac{\partial \theta}{\partial \psi}$, such that elastic storage can be ignored by setting $S_S = 0$.

The ψ form or mixed form of RE with elastic storage included can be shown to work for saturated conditions (Miller et al., 1998) and so can be considered a general governing equation for flow in soils, aquitards, and confined and unconfined aquifers. Some numerical solutions to the ψ form of RE are reportedly subject to poor mass conservation (Milly, 1984, 1985; Celia et al., 1990; Farthing and Ogden, 2017; Clark et al., 2021; Tubini and Rigon, 2022), though it has been shown that mass balance errors can be effectively controlled using ATS (Rathfelder and Abriola, 1994; Tocci et al., 1997). We carefully assess the mass balance performance of our model in this paper.

We consider a finite-difference numerical solution for the ψ form of RE that applies the method of lines to reduce the partial differential equation (PDE) in Eq. (7) to a system of ODEs of the form

$$\frac{d\psi}{dt} = f(\psi, t, z), \quad (8)$$

where ψ and z represent vectors containing discrete values of ψ and z . A similar approach was taken by Tocci et al.

(1997) and Farthing and Ogden (2017). There are various methods that can be applied to integrate Eq. (8) with respect to time – for our purposes, it is useful to consider three specific classes of methods.

- *Fixed-step non-iterative methods (NIMs)*. These methods provide a solution for a fixed time interval and can be either explicit or implicit. Explicit methods are simple but impractical because they require very small time steps. Semi-implicit non-iterative methods take a single iteration based on the linearization of the spatially discretized governing equations. These methods are more stable than explicit methods but have limitations for solving RE over a fixed time interval because they do not iterate to improve the convergence of the problem, which is problematic due to the non-linear dependence of K and C on ψ (Celia et al., 1990). Celia et al. (1990) did not present results using a NIM, but their Eq. (4) can be solved directly (a semi-implicit non-iterative method, which we implemented and discuss in Sect. 3.1). We include NIM here because they may be instructive to new researchers learning about these problems.
- *Fixed-step iterative methods (FIMs)*. Iterative implicit methods are very widely used (Celia et al., 1990; Rathfelder and Abriola, 1994). The iterations allow the solution to find more representative values of K and C over the time step. These methods can be applied to either the ψ form or mixed form of RE. For the mixed form of RE, the convergence criterion can be based directly on water balance closure, as in Celia's modified Picard iteration method. For the ψ form of RE, the convergence criterion is based on ψ values, and the solution is subject to larger water balance errors (Celia et al., 1990; Rathfelder and Abriola, 1994; Farthing and Ogden, 2017).
- *Adaptive time-stepping (ATS) methods*. An implicit or explicit method is used to find $\psi_{t+\Delta t}$ over some *calculation* time step; the truncation error in the solution is assessed, for example, by comparing two different-order solutions, as in Kavetski et al. (2001); and depending on the size of the error, the time step is either reduced to improve the accuracy or increased to improve the efficiency. The solution marches forward until reaching what we term the *reporting* time step, where the state variables are output. The advantage of this approach is that the states and fluxes calculated at the intermediate calculation steps contain useful information that can be exploited in the output, as we demonstrate in this paper. Kavetski et al. (2001, 2002a, b) developed ATS solvers designed specifically to solve the different forms of RE, while other workers have applied readily available “black-box” (Kavetski, et al., 2001) ODE or DAE (differential algebraic equation) solvers to RE (Tocci et al., 1997; Ireson et al., 2009; Ireson and Butler, 2013; Mathias et al., 2015; Clark et al., 2021). A

wide range of ODE solvers are available in many different programming languages. Solutions are simple to implement and, as we show in this study, can outperform other methods in terms of accuracy and efficiency.

In this study, we implement each of the three possible methods for solving RE in scripts that are available from Ireson (2022, <https://github.com/amireson/openRE>, last access: 24 January 2023). All of the models in this study were coded in Python (version 3.8.11). We make use of the following libraries: NumPy (version 1.20.3); Matplotlib (version 3.4.2); SciPy (version 1.7.1), which contains various ODE solvers, described below; and Numba (version 0.53.1, Lam et al., 2015), which is a just-in-time (JIT) compiler that is optional but speeds up the model runs considerably. We organize and run the models using makefiles (Jackson, 2016).

We also provide a MATLAB version of our recommended solution (implemented in MATLAB R2017b installed on a Mac). The MATLAB implementation will not be described further, but compared with the optimal Python solution, the MATLAB model is somewhat simpler and achieves an equivalent performance in terms of simulated states and fluxes. We do not recommend Python over MATLAB (or vice versa) – both platforms work well, and the choice will come down to numerous factors, including, e.g., the availability of, and user familiarity with, either package or the need to use non-proprietary software to satisfy open-science requirements.

2.1 Fixed-step non-iterative and iterative solutions

For the NIM and FIM methods, we have coded up the numerical solutions from Celia et al. (1990). This provides a typical NIM method (first-order backward Euler implicit solution, Eq. 4 in Celia et al., 1990) and two alternative FIMs: Picard iteration that solves the ψ form of RE and their improved modified Picard iteration method (MPM) solution that solves the mixed form of RE. Fully reproducible details of these models were provided by Celia et al. (1990) and so will not be repeated here. These models were implemented in Python, and the code is provided at Ireson (2022, <https://github.com/amireson/openRE>, last access: 24 January 2023). The NIM model was coded up in 67 lines of code (note that blank lines and comment lines are not counted in the number of lines of code), with an additional 52 lines of code to configure the problem (define the grid, hydraulic properties, etc.). The FIM Picard iteration method was coded up in 79 lines of code and required the same 52 lines of code to configure the problem. The FIM modified Picard iteration was implemented using the Numba just-in-time compiler (Sect. A5) and was coded up in 90 lines of code, with an additional 59 lines of code to configure the problem. All solutions make use of the Thomas algorithm to solve the tridiagonal linear system arising from the implicit method.

2.2 Adaptive time-stepping solution

One major benefit of using a standard ODE solver to integrate RE is that the code is concise and easy to read and understand. Our objective is to write a function that will evaluate the right-hand side of Eq. (8) and feed this to the ODE solver. Within this function, the problem can be treated as instantaneous in time so that we only need to consider the spatial differences in our finite-difference solution scheme. We will use a cell-centered grid (Bear and Cheng, 2010, p. 533) in space; i.e., state variables are stored at nodes located at the center of grid cells, while fluxes are defined at the cell boundaries, as shown schematically in Fig. 1 (note that this is equivalent to what is called a “staggered grid” in fluid dynamics). For simplicity here, we consider a uniform grid (constant Δz), but it is straightforward to adapt these solutions to non-uniform grids. We introduce two spatial indices (Fig. 1): i represents the nodal values, and j represents the cell boundaries, both of which have an initial value of zero (because Python uses zero-based indexing). Hence, $j = i + 1/2$. We will start with the ψ form of RE, for which the governing equation is given in the form

$$\frac{\partial \psi}{\partial t} \Big|_i = - \frac{1}{C(\psi_i)} \frac{\partial q}{\partial z} \Big|_i \tag{9}$$

and

$$\frac{\partial q}{\partial z} \Big|_i = \frac{q_{i+\frac{1}{2}} - q_{i-\frac{1}{2}}}{\Delta z} = \frac{q_{j+1} - q_j}{\Delta z}. \tag{10}$$

Here, the fluxes are given by

$$\begin{aligned} q_{i+\frac{1}{2}} = q_{j+1} &= - \frac{K(\psi_{i+1}) + K(\psi_i)}{2} \\ &\times \left(\frac{\psi_{i+1} - \psi_i}{\Delta z} - 1 \right) \\ q_{i-\frac{1}{2}} = q_j &= - \frac{K(\psi_i) + K(\psi_{i-1})}{2} \\ &\times \left(\frac{\psi_i - \psi_{i-1}}{\Delta z} - 1 \right). \end{aligned} \tag{11}$$

In Eq. (11), we are using the arithmetic mean of K at the nodal points to estimate K at the cell boundaries, but other options are possible (see, e.g., Bear and Cheng, 2010, p. 535). It is possible to combine Eqs. (9)–(11), but keeping them separate keeps the code modular and simple.

There are three commonly used boundary conditions for this problem, namely

- i. a specified flux, q_T (LT^{-1}) (type-II boundary) is often used at the upper boundary to represent infiltration, where

$$q_{j=0} = q_T; \tag{12}$$

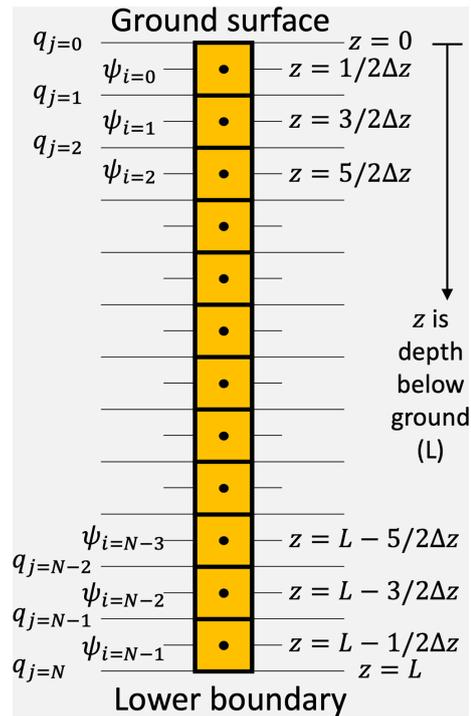


Figure 1. Schematic representation of the cell-centered finite-difference grid, for a soil column of depth L , showing the zero-based Python indices of the states (ψ_i) and fluxes (q_j) on the left and depths of the nodes on the right, assuming a regular grid. There are N states and $N + 1$ fluxes. The upper- and lower-boundary conditions are $q_{j=0}$ and $q_{j=N}$, respectively.

- ii. a free-drainage boundary is often used at the lower boundary, where

$$q_{j=N} = K(\psi_{i=N-1}); \tag{13}$$

- iii. a fixed ψ (type-I) boundary maybe used at the upper boundary, ψ_T (L), typically to indicate a ponding depth, or at the lower boundary, ψ_B (L), typically to represent a fixed water table, where

$$\begin{aligned} q_{j=0} &= - \frac{K(\psi_T) + K(\psi_{i=0})}{2} \\ &\times \left(\frac{\psi_T - \psi_{i=0}}{\Delta z/2} - 1 \right) \\ q_{j=N} &= - \frac{K(\psi_{i=N-1}) + K(\psi_B)}{2} \\ &\times \left(\frac{\psi_{i=N-1} - \psi_B}{\Delta z/2} - 1 \right). \end{aligned} \tag{14}$$

Box 1 provides Python-based pseudo-code that implements this solution (Eqs. 9–14) in a function for $d\psi/dt$ with a type-II boundary at the upper boundary and a free-drainage boundary at the lower boundary and contains just seven lines of code. This function can be called by the ODE solver.

```

# Function inputs: psi,pars,dz,n,qT,psi_old
# psi is an array containing the dependent variable
# pars is a dictionary containing the soil parameters
# dz is the space step
# n is the number of spatial nodes
# qT is the upper boundary flux
# psi_old is psi at the previous time step (needed for
# the mixed-form of RE)

# Function outputs: dpsidt
# dpsidt is an array containing the derivative dpsidt

# Python based pseudo code:
# Kfun function that returns the hydraulic conductivity
# Cfun function that returns the specific storage

# Calculate K and C at each node
K=Kfun(psi,pars)
C=Cfun(psi,pars,psi_old)

# Initialize the flux array for grid boundaries:
q=zeros[n+1]
# Upper boundary condition:
q[0]=qT
# Fluxes for internal nodes:
q[1:-1]=-(K[1:]+K[:-1])/2*((psi[1:]-psi[:-1])/dz-1)
# Lower boundary condition:
q[-1]=K[-1]
# Continuity equation:
dpsidt=-(1/C)*(q[1:]-q[:-1])/dz

```

Box 1. Python-based pseudo-code implementation of RE with a cell-centered finite-difference approach in a function to be called by an ODE solver. Arrays are zero-indexed; $q[0]$ and $q[-1]$ refer to the first and last item in the q array, respectively, and $K[1:]$ and $K[:-1]$ refer to a slice of the array K from the second to the last node and from the first to the second-to-last node, respectively.

2.3 Mass balance closure

When RE is solved over some time interval, $t = t_0 \rightarrow t_M$ (where $t_M - t_0$ typically corresponds to multiple years in practical application) for a soil profile $0 \leq z \leq L$, the cumulative inflow minus outflow should equal the change in storage in the profile over the same interval; i.e., ϵ_B (mm), the bias error, defined in Eq. (15), should be zero.

$$\epsilon_B = \int_{t=t_0}^{t_M} (q(t, 0) - q(t, z_N)) dt - \int_{z=0}^L (\theta(t_M, z) - \theta(t_0, z)) dz \quad (15)$$

The bias error can be treated as a mass balance performance metric for the model, but this metric may underestimate the true errors in the water balance that occur within the time period simulated, which may cancel out over the entire

run. The metric used by Celia et al. (1990), Rathfelder and Abriola (1994), and Tocci et al. (1997) has the same problem. A more rigorous mass balance performance metric is the root mean squared error of the daily (or some other time increment) cumulative net flux minus the change in storage, ϵ_R (mm), where

$$\epsilon_R = \sqrt{\sum_{j=1}^M \left(\begin{array}{l} \int_{t=t_j}^{t_{j+1}} (q(t, 0) - q(t, z_N)) dt \\ - \int_{z=0}^L (\theta(t_{j+1}, z) - \theta(t_j, z)) dz \end{array} \right)^2} / M, \quad (16)$$

where j is an index in time, and M is the number of time steps considered. Reporting both metrics is informative – a high ϵ_R with a low ϵ_B indicates that daily errors are occurring but canceling one another out; a high ϵ_B with a low ϵ_R indicates that small daily errors are systematic in one direction and hence accumulate to give a high bias.

The fluxes in the mass balance calculation depend on the boundary conditions. For type-I (specified ψ) and free-

drainage-type boundary conditions, the boundary flux depends on the simulated ψ value at the node closest to the boundary. Over a time increment t_j to t_{j+1} , ψ values will change continuously and hence so will the boundary flux. Due to the non-linearity of the $K(\psi)$ relationship, the flux will change in a non-linear manner. The cumulative flux over the time increment, $Q_{j \rightarrow j+1}$ (mm), is given by

$$Q_{j \rightarrow j+1} = \int_{t=t_j}^{t_{j+1}} q(t) dt. \tag{17}$$

$Q_{j \rightarrow j+1}$ is estimated from discrete values of q (which in turn are approximated from discrete values of ψ , using either a forward difference approximation, where $Q_{j \rightarrow j+1} \approx q_j \Delta t$; a backward difference approximation (as in Celia et al., 1990), where $Q_{j \rightarrow j+1} \approx q_{j+1} \Delta t$; or a central difference approximation (as in trapezoidal integration), where $Q_{j \rightarrow j+1} \approx (q_j + q_{j+1}) \Delta t / 2$). These discrete approximations for $Q_{j \rightarrow j+1}$ can be poor if the time step is large – or, more precisely, if the changes in q over a time step are large and non-linear.

This leads to an important limitation with Celia's mixed form solution to RE (and other equivalent iterative solutions). This solution has excellent mass balance closure, but because it uses a fixed-step iterative solution procedure with a backward difference approximation for $Q_{j \rightarrow j+1}$, the simulated boundary fluxes can be shown to be sensitive to the time step (see Sect. 3.2). Hence, even though for larger Δt the water balance is still perfectly closed, the actual terms within the water balance have changed, so there is less inflow and less change in storage.

ATS solvers can provide a practical solution to solving RE with good mass balance performance and boundary fluxes that do not depend on the user-specified time step. The basic idea behind ATS solvers is that when there are large changes in ψ in the model, small steps can be taken to capture the shape of $C(\psi)$ and $Q_{j \rightarrow j+1}$ and minimize the integration errors over a time step. When the changes in ψ are small, larger steps can be taken to maximize efficiency. We will refer to these adaptive time steps as *calculation* steps. The user specifies the time steps at which they wish the results to be saved, which we will refer to as the *reporting* time steps. In typical practical applications, the reporting step would be the resolution of the driving data (e.g., hourly or daily). The solver may take many calculation steps of varying lengths between the reporting steps, saving the outcomes internally each time that the error tolerance is satisfied and a successful step is taken.

To accurately calculate the boundary fluxes, it is necessary to use the ψ information from the calculation time steps because ψ may have evolved non-linearly over the reporting time step. However, this is not trivial. Two possible ways to do this (which are equivalent to one another) are to (i) enable dense output from the ODE solver (if this feature is supported

by the ODE solver) or (ii) force the ODE solver to make the reporting steps equal to the calculation steps. However, both of these approaches are more computationally demanding in terms of memory and runtime – a significant disadvantage. We propose here a third method for calculating the boundary fluxes that can be used with any ODE solver. At an instance in time, the cumulative boundary flux, Q , is related to the instantaneous flux, q , by

$$\frac{dQ}{dt} = q. \tag{18}$$

We can therefore use the ODE solver to integrate this expression and solve for Q . To do this, we add to the system of ODEs defined by Eq. (8) two new ODE expressions that represent the cumulative boundary fluxes. The dependent variable vector that is sent to the ODE solver now is F , defined

$$F = \begin{bmatrix} Q_{T:t_0 \rightarrow t} \\ \psi_0 \\ \psi_1 \\ \vdots \\ \psi_{N-2} \\ \psi_{N-1} \\ Q_{B:t_0 \rightarrow t} \end{bmatrix}, \tag{19}$$

where $Q_{T:t_0 \rightarrow t}$ and $Q_{B:t_0 \rightarrow t}$ are the cumulative boundary fluxes at the current time, t , since the start of the simulation, t_0 . The ODE solver will integrate the equation

$$\frac{dF}{dt} = f(F, t, z) \tag{20}$$

to solve for F . The function that is called by the ODE solver will evaluate the vector

$$\frac{dF}{dt} = \begin{bmatrix} \frac{dQ_{T:t_0 \rightarrow t_j}}{dt} \\ \frac{d\psi_0}{dt} \\ \frac{d\psi_1}{dt} \\ \vdots \\ \frac{d\psi_{N-2}}{dt} \\ \frac{d\psi_{N-1}}{dt} \\ \frac{dQ_{B:t_0 \rightarrow t_j}}{dt} \end{bmatrix} = \begin{bmatrix} q_T \\ -\frac{1}{C(\psi_0)} \frac{\partial q}{\partial z} \Big|_0 \\ -\frac{1}{C(\psi_1)} \frac{\partial q}{\partial z} \Big|_1 \\ \vdots \\ -\frac{1}{C(\psi_{N-2})} \frac{\partial q}{\partial z} \Big|_{N-2} \\ -\frac{1}{C(\psi_{N-1})} \frac{\partial q}{\partial z} \Big|_{N-1} \\ q_B \end{bmatrix}. \tag{21}$$

We note that each term in Eq. (21) is expressed at a single instant in time, t , and subscripts 0, 1, ..., $N-2$, $N-1$ refer to the indices of the finite-difference discretization points, and we use zero-based indexing to be consistent with the Python language and Fig. 1. After solving for F , the first and last rows of F correspond to the cumulative boundary fluxes, which at time t are $Q_{T:t_0 \rightarrow t}$ and $Q_{B:t_0 \rightarrow t}$. The cumulative fluxes over each time step are obtained from

$$\begin{aligned} Q_{T:t_j \rightarrow t_{j+1}} &= Q_{T:t_0 \rightarrow t_{j+1}} - Q_{T:t_0 \rightarrow t_j} \\ Q_{B:t_j \rightarrow t_{j+1}} &= Q_{B:t_0 \rightarrow t_{j+1}} - Q_{B:t_0 \rightarrow t_j}. \end{aligned} \tag{22}$$

Solving systems of ODEs in this way is straightforward – requiring the user to pack and unpack the dependent variable vector. Python-based pseudo-code showing how this can be implemented is given in Box 2. Note that here we multiply the fluxes by 1000 within the solver (equivalent to converting the fluxes from m d^{-1} to mm d^{-1}), so that the magnitude of the fluxes is comparable with the magnitude of the changes in ψ (typically in m), which can improve the water balance estimate. Any arbitrary scaling factor can be applied here, as long as the output fluxes are re-scaled to the correct units. We shall refer to this method for calculating the boundary fluxes as the *solver flux output method* (SFOM).

2.4 Improving efficiency

Here, we provide details of two techniques that can be used to improve the computational runtime of the model. These methods have no impact on the accuracy of the solution, so they are optional, but combined they can result in better than a factor of 10 reduction in the runtime at the cost of only a few additional lines of code. In Appendix A, we investigate the impact of a range of different possible model decisions/assumptions on the accuracy, mass balance, efficiency, and simplicity of the model.

2.4.1 Defining the Jacobian pattern

For the form of RE given by Eq. (9), the Jacobian matrix, \mathbf{J} , is an $n \times n$ matrix, where the cell in each row, i , and column, j , is defined by the derivative

$$J_{i,j} = \frac{\text{d}}{\text{d}\psi_j} \left(\frac{\text{d}\psi}{\text{d}t} \Big|_i \right) = \frac{\text{d}}{\text{d}\psi_j} \left(-\frac{1}{C(\psi_i)} \frac{\partial q}{\partial z} \Big|_i \right). \quad (23)$$

Each entry in $J_{i,j}$ can be evaluated in a function that is passed to the ODE solver (assuming that the particular ODE solver being used has this functionality), in order to speed up the solution process. For the spatial discretization scheme described above, all the terms of the Jacobian are zero, except for where $j = i - 1$, $j = i$, and $j = i + 1$ (ignoring when $i - 1 < 1$ and $i + 1 > n$, which would be zero terms anyway, for any boundary condition). A simpler alternative to defining the full Jacobian matrix is to define the Jacobian pattern. The Jacobian pattern is a matrix of ones and zeros that defines the location of the structurally non-zero elements of the Jacobian – that is, where the terms are not identically zero. For the spatially discretized RE as given in Eq. (9), the Jacobian pattern is a simple tridiagonal matrix, with ones on the three main diagonals and zeros everywhere else. To implement this requires an ODE solver capable of using the Jacobian pattern (also referred to as the Jacobian sparsity matrix). The SciPy ODE solvers *ode* and *solve_ivp* have the ability to define a banded Jacobian pattern: setting *uband* and *lband* arguments to 1 tells the ODE solver that the Jacobian is a tridiagonal matrix. The SciPy ODE solver *solve_ivp* can also handle a general $n \times n$ Jacobian pattern, which is more adaptable for multi-dependent variable coupled problems (e.g.,

Goudarzi et al., 2016). The MATLAB ODE solvers can read the Jacobian sparsity pattern matrix from the *JPattern* argument. We report on the relative performance/complexity of each of these methods in Sect. A4.

2.4.2 Just-in-time compilation

In this paper, we are providing guidance for the use of interpreted programming languages (e.g., Python or MATLAB) to solve RE. Interpreted languages have a number of advantages over compiled languages (such as FORTRAN, C, and C++), including that they are, at least in our opinion, easier to learn, with excellent teaching resources widely and freely available; they tend to have higher level abstractions, so that the same task can be completed in fewer lines of code; and they are cross platform and typically easier to install. Interpreted languages are not pre-compiled and are hence slower to execute than compiled languages. A nice compromise between the simplicity of interpreted languages and efficiency of compiled languages is to use a just-in-time compiler. In Python, the Numba library is such a just-in-time compiler (Lam et al., 2015). Numba compiles selected Python functions once at the start of the runtime, and then all subsequent calls to the code run much faster. We find that using Numba in conjunction with our preferred ODE solver solution described above, results in up to $10\times$ faster code execution (see Sect. A5). The drawback to using Numba is that some re-factoring of the code may be necessary to make a script that previously ran without Numba work using Numba – in particular, there are complications around how variables are allocated into NumPy arrays. We include code in Ireson (2022, <https://github.com/amireson/openRE>, last access: 24 January 2023) that demonstrates how to successfully implement Numba.

3 Benchmarking the model performance

In this section, we run our RE solver, openRE, on a number of benchmark problems, comparing the different solution procedures and assessing the performance of all solutions against the five success criteria identified in the introduction, namely (i) accuracy of $\theta(t, z)$ and $\psi(t, z)$, (ii) mass balance performance, (iii) consistent boundary fluxes with Δt , (iv) computational efficiency (i.e., runtime), and (v) simplicity of the code. For the purposes of comparing efficiency (iv), all simulations were run on the same laptop computer, and we report the runtimes as a measure of relative performance. For the purposes of comparing simplicity of the code (v) we use a very simple metric of lines of code, which is reported above in Sect. 2.

```

# ODE function inputs: tSpan,F,args
# tSpan array with the start and end date for the integration
# F array with the dependent variable
# args list of other variables (see Box 1)

# ODE function outputs: dFdt
# dFdt array with the time derivative of the dependent variable

# Python based pseudo code for the ODE function:
# Unpack the dependent variable:
QT=F[0]/1000.      # Cumulative flux at the upper boundary
QB=F[-1]/1000.    # Cumulative flux at the lower boundary
psi=F[1:-1].      # matric potential at each node

# Insert here the contents of the function in Box 1, which
# calculates dpsidt, q[0] and q[-1]

# Pack up the dependent variable in a single column array:
dFdt=hstack([q[0]*1000.,dpsidt,q[-1]*1000.])

# Pseudo code to call the ODE solver:
# Initialize the 2D array (t,z) for the dependent variable
F=zeros(nt,nz)

# Pack up the initial condition:
F[0,0]=0.         # Cumulative flux at the upper boundary
F[0,1:-1]=psiI
F[0,-1]=0.       # Cumulative flux at the lower boundary

# Solve using a time loop:
for j in 1 to M-1:
    F[j+1,:]=CallODE(ODEfunction,tSpan=[0,dt],F[j,:],args)

# Unpack dependent variable:
QT=hstack([0,F[1:,0]-F[:-1,0]])/1000.
psi=F[:,1:-1]
QB=hstack([0,F[1:,-1]-F[:-1,-1]])/1000.

```

Box 2. Python-based pseudo-code implementation of an ODE solver solution. The ODE function returns the time derivative of the dependent variable array, given in Eq. (21). Note that $F[:, 1 : -1]$ is a slice through the array F that takes all the items in the first (time) dimension and the second to second-to-last items in the second dimension.

3.1 Published model benchmarks

3.1.1 Celia's problem

Celia's test case (Celia et al., 1990) is used to compare our ATS solution with the different solutions schemes previously proposed by Celia et al. (1990). The test problem uses a 40 cm deep vertical soil profile ($z_N = 40$) with a uniform 1 cm space step ($dz = 1$), a 360 s duration ($t_0 = 0; t_M = 360$) with a 1 s time step ($\Delta t = 1$), and the following initial and type-I boundary conditions:

$$\begin{aligned}
 \psi(t = t_0, 0 \leq z \leq z_N) &= -61.5 \text{ cm} \\
 \psi(t_0 \leq t \leq t_M, z = 0) &= -20.7 \text{ cm} \\
 \psi(t_0 \leq t \leq t_M, z = z_N) &= -61.5 \text{ cm.}
 \end{aligned} \tag{24}$$

The soil hydraulic properties are given by

$$\begin{aligned}
 \theta &= \frac{\alpha(\theta_s - \theta_r)}{\alpha + |\psi|^\beta} + \theta_r \\
 K &= K_s \frac{A}{A + |\psi|^\gamma},
 \end{aligned} \tag{25}$$

where the parameter values are $\alpha = 1.611 \times 10^6$, $\theta_s = 0.287$, $\theta_r = 0.075$, $\beta = 3.96$, $K_s = 0.0944 \text{ cm s}^{-1}$, $A = 1.175 \times 10^6$, and $\gamma = 4.74$. Celia et al. (1990) presented three solution schemes: the “no-iteration scheme” uses the ψ form of RE and solves the problem with a single backward implicit step and no iteration (which we achieved using the Thomas algorithm), the “Picard iteration scheme” also solves the ψ form of RE but uses the Picard iteration method to improve the solution, with errors in ψ used as a convergence criterion, and finally the “modified Picard iteration method” (MPM)

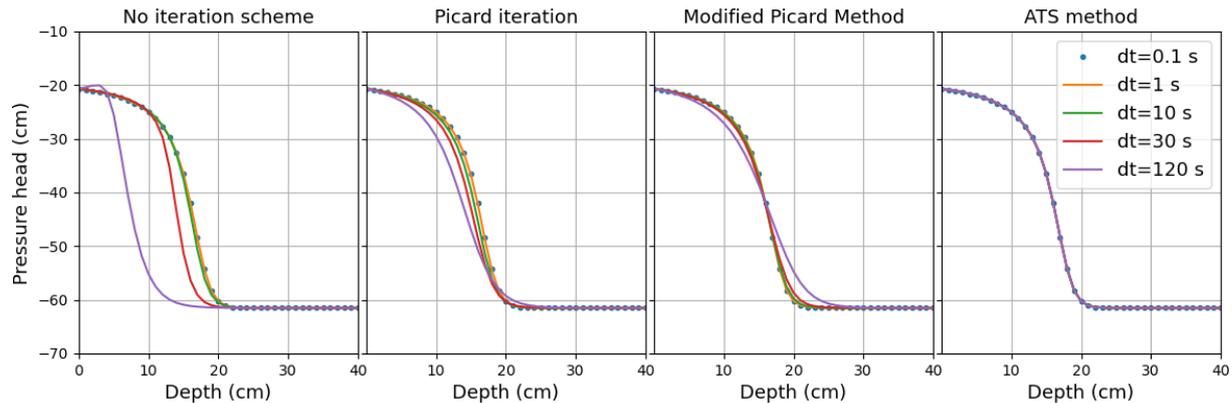


Figure 2. Reproduction of Celia's model benchmark: the ψ breakthrough curve using a backward difference implicit Euler solution with the no-iteration scheme, backward difference implicit Euler with Picard iteration, Celia's modified Picard iteration method (MPM), and the ODE solver with adaptive time stepping. The ODE solver produces consistent $\psi(tz)$ results independently of the mass balance calculation procedure. The time steps reported are calculation time steps for Celia's solutions but are reporting time steps for the ODE solver, which uses an adaptive time step for calculation steps.

uses the mixed-form of RE and uses errors in θ as a convergence criterion. The MPM is mass conservative because the iteration ensures that the cumulative change in fluxes (the right-hand side of RE) balances the changes in storage (the left-hand side of the mixed form of RE). Celia's three solutions were reproduced in Python scripts (<https://github.com/amireson/openRE>, last access: 24 January 2023) and compared with our ATS/SFOM solution. Results from all three solutions are shown in Fig. 2. All methods are consistent for very small time steps. The fixed-step method with no iteration performs poorly, with delayed breakthrough of the wetting front when the time step is large. The solution is improved by using the Picard iteration, but there are still some delays. The MPM has a much better performance but, as with all implicit Euler time-stepping schemes, is still subject to some numerical dispersion for larger time steps (van Genuchten and Gray, 1978). The ATS solution reproduces the ψ breakthrough curve but with no dispersion and no differences associated with the time step. We note that the time steps for plotting the ATS solution only represent the reporting time step – the underlying calculation time steps are likely much smaller (Sect. 2.2).

In Fig. 3, we show the cumulative inflow simulated by each of these models for Celia's benchmark problem, along with the mass balance bias error, for different reporting time steps. The fixed-step solution with no iteration and the Picard iteration solution both have poor mass balance performance unless the time step is very small – on this basis, we do not consider these solutions further. We see that the MPM method is perfectly mass conservative for any Δt used in the model, as we should expect. However, we can see that the cumulative inflow is sensitive to Δt . Hence, even though for larger Δt the water balance is still perfectly closed, the actual terms within the water balance have changed, so there is less inflow and less change in storage. This is perhaps an under-

appreciated limitation of Celia's MPM solution and solutions to the mixed form of RE generally – which is that mass balance is a necessary but insufficient criterion for model performance assessment, and truncation errors can still be present in the fluxes even with perfect water balance closure.

In Fig. 3, we also show the water balance performance of our ATS solution, using either reporting-time-step information for the water balance calculation or using calculation-time-step information (i.e., using the SFOM described in Sect. 2.3). Using reporting-time-step information is the easiest and most intuitive approach to take – you numerically integrate (1) discrete θ values over depth to get storage and (2) discrete q values during reporting time steps to get Q (e.g., using trapezoidal integration). However, this approach fails to capture non-linear changes in q over a reporting time step and results in large water balance errors and errors in the cumulative fluxes, as is clear in Fig. 3. Using the SFOM, we see that the water balance is almost exactly closed, and the boundary fluxes are independent of the reporting time step. It is also important to note that the discrete $\psi(tz)$ values simulated by both ATS solution procedures here are identical (see Fig. 2) – the only difference is how the boundary fluxes are calculated.

3.1.2 Miller's saturated infiltration pulse problem

Miller et al. (1998) investigated solutions to RE that aimed to address numerical convergence problems associated with challenging boundary conditions, and they specifically looked at the problem of infiltration from a ponded upper boundary into a hydrostatic soil profile with a fixed water table at the lower boundary. This is a good benchmark because it requires the model to deal with perched saturated conditions over unsaturated conditions and involves highly non-linear changes in properties over short distances

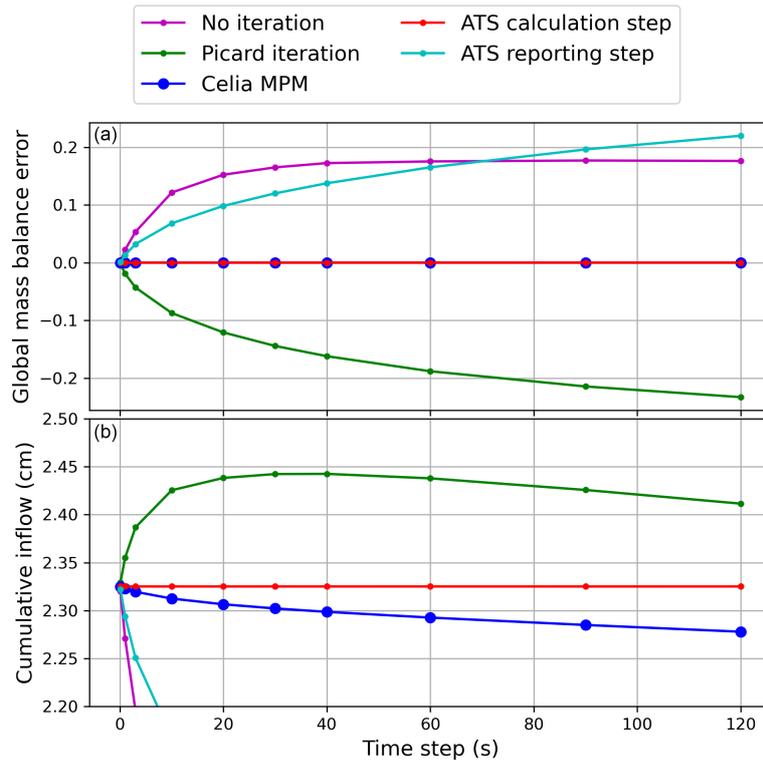


Figure 3. Mass balance of Celia's three models and our adaptive time-stepping model. Panel (a) shows a reproduction of Celia's mass balance calculation. Panel (b) shows the cumulative inflow that is simulated.

and time steps. The problem uses van Genuchten (1980) hydraulic properties, given by

$$S_e = (1 + (\alpha\psi)^n)^{-m}, \quad (26)$$

$$\theta = \theta_r + (\theta_s - \theta_r)S_e, \quad (27)$$

$$\frac{d\theta}{d\psi} = \frac{-\alpha m (\theta_s - \theta_r)}{1 - m} S_e^{1/m} (1 - S_e^{1/m})^m, \quad (28)$$

$$K = K_s S_e^{1/2} (1 - (1 - S_e^{1/m})^m), \quad (29)$$

where S_e (–) is the effective saturation; α (L^{-1}), n (–), and m (–) are parameters that determine the shape of the $\theta(\psi)$ curve; θ_r (–) and θ_s (–) are the residual and saturated volumetric water contents; and K_s (LT^{-1}) is the saturated hydraulic conductivity. Miller's problem uses the parameters in Table 1 for three different soil types.

A hydrostatic initial condition is used, with a fixed water table at depth of 10, 5, and 2 m below ground surface for sand, loam, and clay loam, respectively. At the upper boundary, 0.1 m of ponding is applied throughout the simulation runtime of 0.18, 2.25, and 1.0 d for sand, loam, and clay loam, respectively. We simulated this problem with our ATS solution and with Celia's MPM model for comparison purposes. Both models faced challenges with this problem. For Celia's MPM, we had to use a small time step to get the solver to produce accurate $\psi(z)$ profiles (Fig. 4). For the ATS solutions using the default ODE solver settings, the models

Table 1. Soil hydraulic properties used the Miller et al. (1998) problem.

Parameter	Sand	Loam	Clay loam
θ_r	0.093	0.078	0.095
θ_s	0.301	0.430	0.410
α	5.470	3.600	1.900
n	4.264	1.560	1.310
K_s (md^{-1})	5.040	0.250	0.062
S_S	10^{-6}	10^{-6}	10^{-6}

failed to propagate the wetting front into the soil correctly. It was necessary to increase the maximum number of calculation steps allowed per reporting time step (we increased this from the default 500 to 10 000) so that very small time steps could be taken (Table 2). It was also necessary (for loam and clay loam) or beneficial (for sand) to reduce the solver error tolerances – see values in Fig. 4 and Table 2. The results from these simulations are shown in Fig. 4 and are consistent with those reported in Fig. 1 of Miller et al. (1998), showing that both models are able to successfully reproduce this benchmark. The runtimes and water balance for each solution are tabulated in Table 2. Celia's MPM has consistently better water balance performance, though we think the water balance errors in both models are acceptably low. The ATS solution

Table 2. Runtime and water balance performance for the Celia MPM and ATS SFOM solutions, applied to the problem in Miller et al. (1998). We also show here the solver settings. For the ATS solutions, we had to reduce the relative and absolute tolerances (atol/rtol) of the integrator from the default settings and increase the maximum number of calculations steps allowed within a reporting step (nsteps). Note that time steps (dt) in ATS solutions are reporting time steps.

Solution	Soil	Solver settings	Runtime (s)	MB bias (mm)	MB RMSE (mm)
Celia MPM	Sand	$dt = 0.001$ d	4.4	-4.2×10^{-6}	4.5×10^{-6}
	Loam	$dt = 0.001$ d	31.7	-1.6×10^{-3}	5.3×10^{-6}
	Clay loam	$dt = 0.001$ d	32.6	-8.0×10^{-2}	2.5×10^{-4}
ATS SFOM	Sand	atol/rtol = 10^{-6} nsteps = 10 000 $dt = 0.01$	11.7	-1.5×10^{-2}	8.8×10^{-4}
	Loam	atol/rtol = 10^{-6} nsteps = 10 000 $dt = 0.01$	16.7	-8.8×10^{-2}	2.0×10^{-3}
	Clay loam	atol/rtol = 10^{-5} nsteps = 10 000 $dt = 0.001$	30.8	-3.2×10^{-1}	1.3×10^{-3}

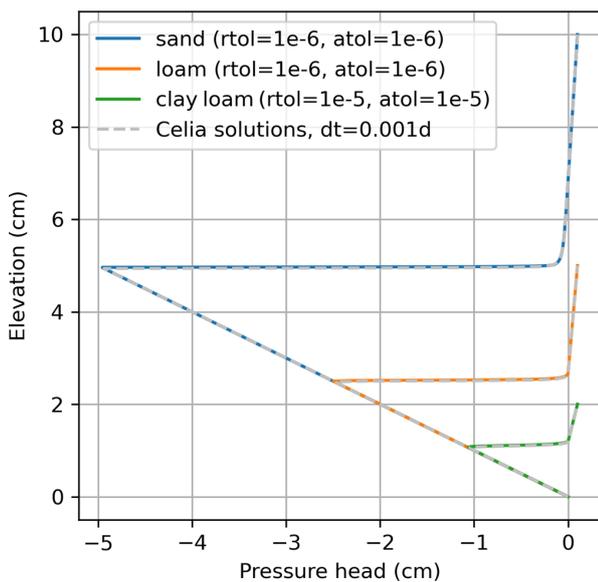


Figure 4. Reproduction of the Miller infiltration pulse result, using Celia's MPM model and our ATS SFOM solution. Both models are satisfactorily consistent with the output reported by Miller et al. (1998), in their Fig. 1.

is slower for sand, faster for loam, and about the same for clay loam. We note that the runtimes and water balances of the ATS solution are sensitive to the reporting time step Δt and the solver settings nsteps, atol, and rtol – an improved solution might be attainable by optimizing these settings. On the other hand, for Celia's MPM solution, we only needed to optimize Δt .

3.1.3 Mathias' solution for horizontal infiltration

Mathias and Sander (2021) developed a pseudospectral similarity solution for horizontal infiltration (i.e., solving RE without gravity) that is fast and accurate. This solution assumes a semi-infinite horizontal soil column ($0 \leq x < \infty$) with a uniform initial condition ($\psi(t = 0) = \psi_0$) and a type-I boundary condition on the left boundary ($\psi(x = 0) = \psi_L$). The model was run for 100 min ($0 \leq t \leq 100$ min). The solution can resolve very large gradients in saturation and ψ at the boundary that propagate into the soil rapidly – and as such this is another challenging problem for a numerical RE model to reproduce. We solved this problem for three soil types, namely, Hygiene sandstone, silt loam GE 3, and Beit Netofa clay, with properties from van Genuchten (1980) as listed in Table 3. S_s was set to 0, consistent with Mathias' solution. To configure our model for horizontal flow, it is necessary to remove the gravity term from the flux calculations (i.e., Eq. 11). We solved this problem for a left-hand boundary effective saturation of 0.99 and an initial saturation of 0.01. The grid is configured such that the wetting pulse does not reach the right-hand boundary over the simulation runtime.

We solved this problem with our ATS solution, with Celia's MPM solution, and with the pseudospectral similarity solution (Mathias and Sander, 2021, implemented in MATLAB). The results in Fig. 5 show that both the ATS solution and the Celia solution do an excellent job of reproducing this solution for $\theta(tx)$ (where x , m, is horizontal distance). The runtimes and water balance for each solution are tabulated in Table 4. Here, we see that the ATS and Celia solutions have the same performance in terms of the water balance and the cumulative fluxes simulated. Runtimes vary between models:

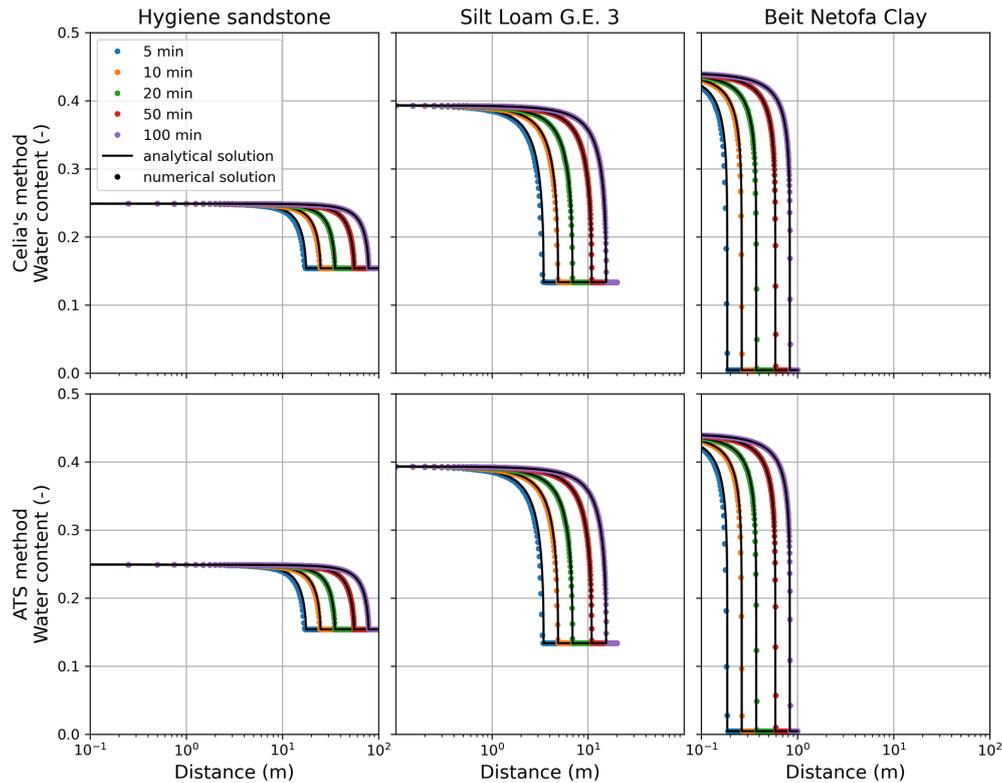


Figure 5. Comparison of our ATS solver flux output model for horizontal infiltration with the Mathias and Sander (2021) pseudospectral similarity solution (denoted analytical in the legend).

Table 3. Soil hydraulic properties used the Miller et al. (1998) problem.

Parameter	Sand	Loam	Clay loam
θ_r	0.153	0.131	0.
θ_s	0.250	0.396	0.446
α	0.0079	0.00423	0.00152
n	10.4	2.06	1.17
K_s (cm d ⁻¹)	108.	4.96	0.082
S_S	0.	0.	0.
dt (min)	0.1	0.01	0.01
dx (cm)	0.25	0.05	0.0025

both are the same for sandstone, Celia's solution is faster for silt loam, and the ATS solution is faster for clay.

3.2 Comparison with Hydrus 1D

Hydrus 1D (Šimůnek et al., 2005, 2016) is a widely used one-dimensional RE solver. The calculations within Hydrus are undertaken using openly available FORTRAN source code, and the software runs through a (closed-source) graphical user interface on Microsoft Windows. The FORTRAN code can be compiled using gfortran on the macOS operating system and run from the command line, which we did

here, so that the runtime comparisons with our model are fair. Within Hydrus, the user interface provides somewhat limited control over the error tolerances. We were unable to modify any settings to improve the water balance, and so we present here model runs that use the default iteration criteria (maximum number of iterations = 100; water content tolerance = 0.001; pressure head tolerance = 10 mm; lower/upper optimal iteration range = 0.7/1.3; lower/upper time step multiplication factor = 1.3/0.7; lower/upper limit of tension interval = $10^{-6}/10^3$ cm).

We configured Hydrus 1D, our ATS solution, and our implementation of Celia's MPM solution for a simple numerical experiment, where we simulate the infiltration of a 10-year time series of daily precipitation into a 1.5 m deep soil column, with a free-drainage lower-boundary condition. The minimum, mean, and maximum annual precipitation was 265, 484, and 680 mm yr⁻¹, and the maximum daily precipitation was 55 mm d⁻¹. We used silt loam GE 3 soil hydraulic properties from van Genuchten (1980), where $\theta_r = 0.131$, $\theta_s = 0.396$, $\alpha = 0.423$ m⁻¹, $K_s = 0.0496$ m d⁻¹, and $n = 2.06$. We set $S_S = 10^{-6}$ m⁻¹ and used a uniform ψ initial condition of -3.59 m. The results of the simulations with each model are given in Table 5, showing the runtime and water balance performance; Fig. 6, showing the detailed water balance performance; and Fig. 7, showing the simulated storage and drainage.

Table 4. Runtime and water balance performance for the Celia MPM and ATS SFOM solutions, applied to the problem described by Mathias and Sander (2021).

Solution	Soil	Runtime (s)	MB bias (mm)	MB RMSE (mm)	Cumulative infiltration (mm)
Celia MPM	Sandstone	2.4	4.0×10^{-3}	6.5×10^{-6}	63.2
	Silt loam	12.8	2.2×10^{-2}	3.8×10^{-6}	34.2
	Clay	69.4	2.2×10^{-3}	4.1×10^{-7}	3.4
ATS SFOM	Sandstone	2.3	-1.1×10^{-3}	4.8×10^{-6}	63.3
	Silt loam	18.8	1.9×10^{-2}	3.9×10^{-6}	34.2
	Clay	32.1	2.2×10^{-3}	3.9×10^{-7}	3.4

Table 5. Comparison of model runtimes and mass balance performance.

Model	Runtime (s)	MB bias (mm)	MB RMSE (mm)
ATS solution, $\text{rtol} = 10^{-6}$	1.71	-0.018	8.06×10^{-5}
ATS solution, $\text{rtol} = 10^{-7}$	2.30	0.0003	6.92×10^{-5}
Celia MPM solution	3.2	0.0	2.3×10^{-10}
Hydrus	2.21	-0.0021	4.05×10^{-3}

In our ATS solution, we can trade off between water balance error and the runtime by modifying the `rtol` argument for the ODE solver. We found that the default `rtol` of 10^{-6} had the fastest runtime, but the water balance performance, whilst good enough for all practical purposes, was the worst overall (Table 5, Fig. 6). Therefore, we reduced `rtol` to 10^{-7} , which improved the water balance performance but increased the runtime. Even though the water balance errors reported here are all very small, it is still interesting to look closely at how these compare for the different models, as shown in Fig. 6. The first thing to note is that the Celia MPM solution has water balance errors of essentially zero, which we expect, because this solution enforces water balance closure. Celia's solution did have the longest runtime – approximately 40 % slower than the other solutions. In the ATS solutions, on a daily basis, the water balance errors are much smaller than in Hydrus. However, in Hydrus, the water balance errors appear random, with a mean of zero, and hence when looking at the cumulative errors in Hydrus, there is no systematic accumulation in the errors. In the ATS solutions, in the lower-right panel of Fig. 6, we can see that the water balance errors are strongly correlated to the infiltration flux at the upper boundary – larger fluxes result in larger errors. Hence, for the ATS solution with $\text{rtol} = 10^{-6}$, we see that the errors accumulate, and after 4 years, the cumulative errors in the ATS solution exceed those in Hydrus. For the ATS solution with $\text{rtol} = 10^{-7}$, the errors do not accumulate monotonically, and the long-term cumulative errors tend to oscillate about zero. The water balance performance of the ATS solution with $\text{rtol} = 10^{-7}$ is therefore better than the performance in Hydrus (Table 5, Fig. 6), while the runtimes of these models are essentially the same (Hydrus is slightly faster, with a runtime of 2.21 vs. 2.30 s, Table 5).

Looking at the simulated storage and discharge in Fig. 7, the two ATS solutions are visually indistinguishable and are both broadly consistent with the Hydrus 1D model outputs. The Celia MPM solution has non-negligible differences with all other solutions. This is because the MPM solution applies an iterative solution procedure to solve the model at a daily time step, and the boundary fluxes are therefore subject to errors, as discussed above. The solution scheme imposes mass balance on the problem but does not track the truncation errors in the fluxes. The avoidance of this issue represents a significant advantage of adaptive time-stepping solutions.

4 Summary and recommendations

We developed a simple adaptive time-stepping scheme (ATS) for RE using the interpreted language Python and making use of the SciPy ODE solver *ode*. We also developed a new solver flux output method (SFOM) whereby cumulative boundary fluxes can be included within the dependent variable vector, allowing the determination of highly accurate integrated fluxes over designated time periods. The SFOM is particularly useful for providing reliable assessment of mass balance closure. In principle, SFOM can be implemented in any ODE solver because it does not require any special output (such as dense output) to be available. Our model was coded up in Python and released with the name *openRE* (Ireson, 2022). Our model performed well against our five success criteria: (i) we successfully reproduced benchmark solutions for $\psi(t, z)$ and $\theta(tz)$ from Celia et al. (1990), Miller et al. (1998), and Mathias and Sander (2021); (ii) we report negligibly low mass balance errors; (iii) we simulate boundary fluxes that are independent of the reporting time step (unlike

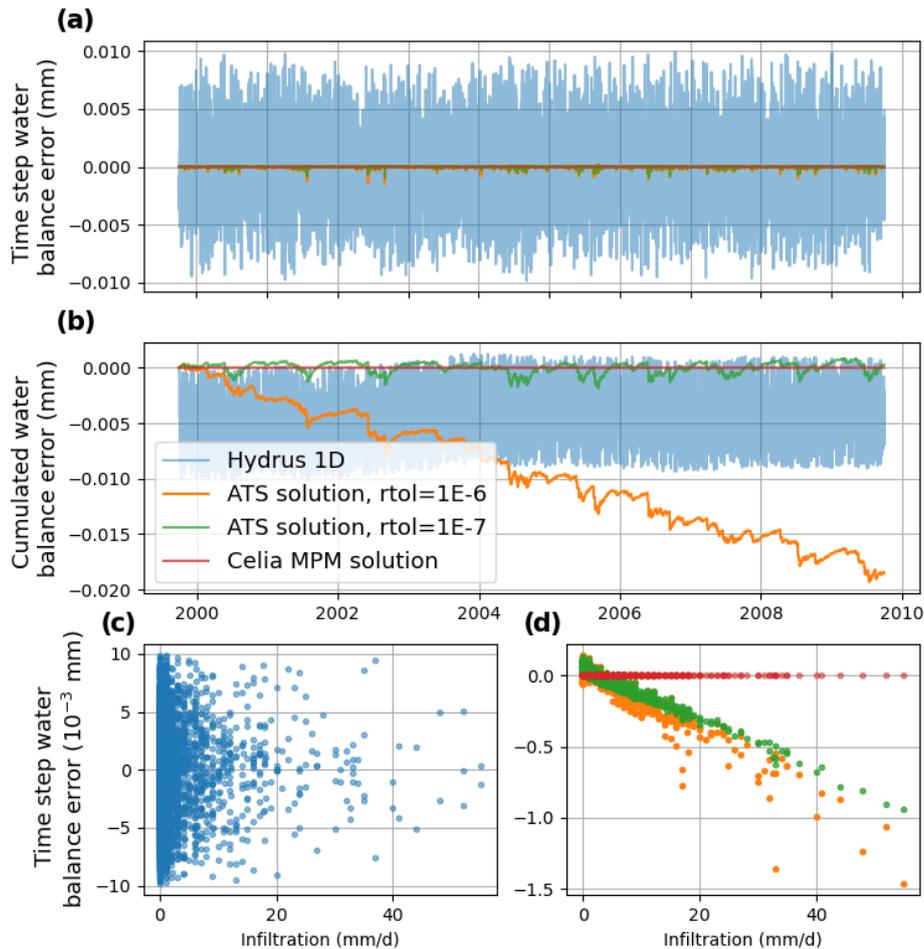


Figure 6. Comparison of water balance performance from Hydrus 1D, the ATS solution, and our implementation of Celia's MPM solution. The water balance error is $q_T - q_B - \Delta S$, and we show the balance for each time step (a) and cumulative balances since the start of the simulation (b). In panels (c) and (d) the time step error is plotted against the infiltration rate.

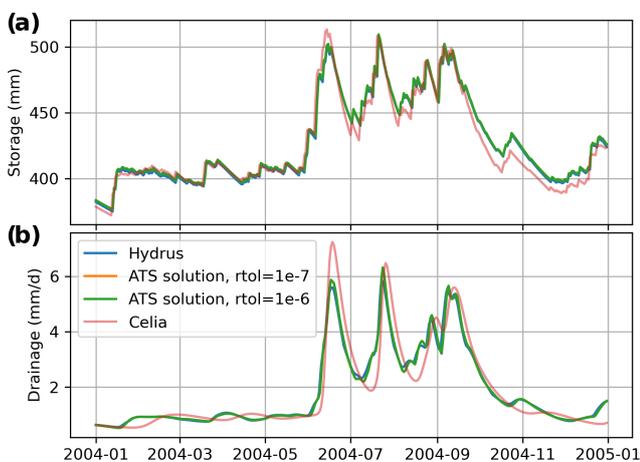


Figure 7. Simulated storage (a) and drainage (b) using Hydrus 1D, the ATS solution, and our implementation of Celia's MPM solution.

Celia's solution, as demonstrated in Fig. 7); (iv) we have low runtimes (as good as Hydrus 1D); and (v) our code is very simple, concise (92 lines of code for the solver plus 68 lines of code for model configuration for the numerical experiment in Sect. 3.2), and easily adaptable to new problems. Our solution had the best balance of efficiency, accuracy, and simplicity as compared to alternative established solution procedures.

Appendix A: Navigating pitfalls in ODE solver solutions

There are several subtle decisions that must be made when solving RE using a generic ODE solver. Here, we test a number of alternative model configurations and report the impact of these decisions using the following metrics: for model accuracy (criteria i), we report the RMSE of ψ at all grid points in t and z between the current model run and a reference model run; for the mass balance (criteria ii), we report both the bias error (Eq. 15) and the more rigorous daily water bal-

ance RMSE (Eq. 16); for the model efficiency (criteria iv), we simply report the runtime, where all runs were undertaken on the same laptop computer. For these numerical experiments, we used the 10-year infiltration numerical experiment described in Sect. 3.2.

The best model configuration, against which all other model configurations are compared, was as follows: use the SciPy ODE solver *ode* with the method BDF (backward differentiation formula, Brown et al., 1989), use our SFOM solution (Sect. 2.3/A2), use the analytical expression for $C(\psi)$, use a banded Jacobian sparsity pattern matrix (Sect. 2.4.1/A4), and use the Numba JIT compiler (Sect. 2.4.2/A5). The water balance performance of this model, showing the cumulative change in storage against cumulative inflow (as infiltration at the surface) minus outflow (as drainage at the base), is plotted in Fig. A1.

A1 Alternative SciPy ODE solvers

Here, we compare the alternative ODE solvers that were available in SciPy at the time of writing, which includes *ode*, *odeint*, and *solve_ivp*. These functions are alternative wrappers to classic ODE solvers written in Fortran, of which we consider here VODE with the method BDF (Brown et al., 1989, available within *ode* and *solve_ivp*) and LSODA (available with all three functions, Petzold, 1983). Note that for all solutions reported here we used the banded Jacobian sparsity pattern, with the exception of *solve_ivp* BDF, which only allows for the full Jacobian sparsity pattern to be defined and which we found slowed the solution down – hence the results for the *solve_ivp* BDF model do not use any information about the Jacobian matrix.

We see that *solve_ivp* underperforms in accuracy, water balance, and efficiency. The *odeint* solver has the best performance in terms of accuracy and water balance but is slower by a non-negligible amount. The *ode* BDF method is the most efficient but has slightly worse water balance performance – however, the water balance performance of all methods is extremely good, and errors are negligible for practical purposes. We therefore chose *ode* BDF as our preferred solution – but *ode* LSODA is also a good option. It is also possible to increase the error tolerances in the ODE solver, reduce the maximum number of time steps, and increase the minimum time step – all of which could result in a faster runtime at the cost of lower accuracy/water balance closure.

A2 Alternative boundary flux calculation methods

As detailed in Sect. 2.3, there are alternative ways to calculate the boundary fluxes for use in the water balance calculation. In Sect. 3.3, we developed a novel approach to calculating the boundary fluxes – the SFOM. In addition to this method, we consider methods that calculate the boundary fluxes based on the output model states at either reporting-step or calculation-step information. We also consider using

forward, backward, or central difference approximations to integrate the flux over a time step (Eq. 18). The results of this analysis are provided in Table A2.

Model state variables are unaffected by the different boundary flux calculation methods. The SFOM has the best water balance performance, both in terms of bias and RMSE. Using calculation-step-level information results in good water balance closure, with the central difference approximation giving the lowest errors. However, the efficiency of this is poor, with runtimes increased by more than a factor of 3. This is because many additional calculations need to be performed outside the ODE solver for each calculation time step. By default, the ODE solver allows up to 500 calculation time steps for every reporting time step – so this is very inefficient. Calculating the boundary fluxes using reporting-time-step information is very efficient and slightly faster than our method, but the water balance errors are significantly larger. These reporting-step errors will increase with an increased reporting time step, as is shown in Fig. 3. Overall then, the SFOM provides the performance of using calculation-step information without the loss of computational efficiency.

The key take-home point here is that the easiest and most obvious approach to calculating the boundary fluxes is to use reporting-step information. This is a bad idea – the mass balance errors are large, and if this is combined with other bad decisions (such as using discrete approximations for $C(\psi)$ as discussed in the next section), the results can be catastrophic (water balance errors > 100 mm).

A3 Alternative estimation methods for $d\theta/d\psi$

When we use a parametric expression for $\theta(\psi)$, such as the van Genuchten equations (Eqs. 26–29), we can obtain an analytical expression for $d\theta/d\psi$, as in Eq. (28), and this can be used to calculate $C(\psi)$ as implemented in RE in Eq. (5). However, depending on the numerical solution procedure that is adopted, this can lead to errors with mass conservation, and it is recommended by some researchers (Rathfelder and Abriola, 1994; Clark et al., 2021) that a discrete approximation is used for $d\theta/d\psi$, whereby

$$\frac{d\theta}{d\psi} = \frac{\theta_n - \theta_{n-1}}{\psi_n - \psi_{n-1}}, \quad (\text{A1})$$

where here n is a time index. This approach could be seen as equivalent to solving the mixed form of RE and can minimize water balance errors in the model that arise because the changes in $d\theta/d\psi$ over a time step are non-linear, as shown by Celia et al., (1990). However, it is necessary to apply this very carefully in the context of ATS methods. The values of θ_{n-1} and ψ_{n-1} must be available from the previous calculation time step and not the previous reporting time step. If reporting-time-step information is used, the model will fail badly because as the calculation steps move forward in time over a reporting step, $C(\psi)$ is constantly referenced back to the beginning of the reporting step. This is clearly

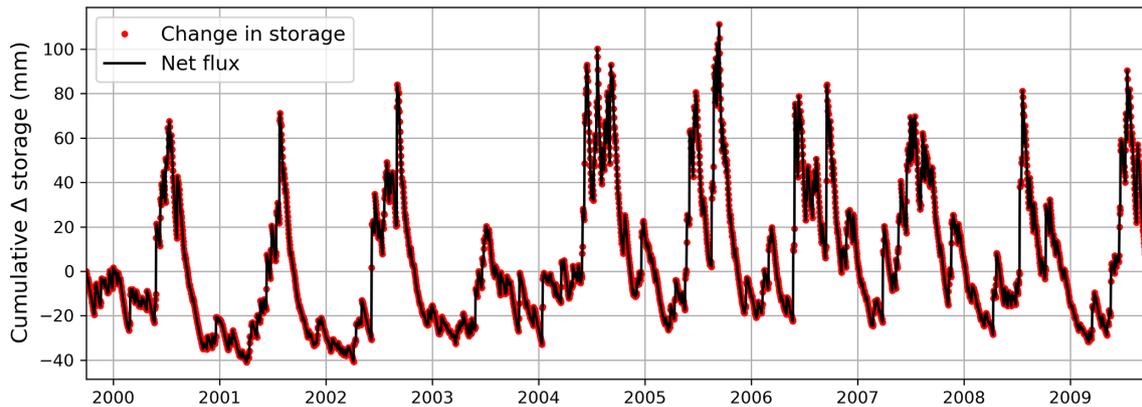


Figure A1. Water balance performance plot for 10-year infiltration experiment, with the best model configuration, showing the cumulative change in storage in the profile and the cumulative inflow minus outflow. The water balance bias was -0.018 mm, and the RMSE of the daily water balance errors was 8.06×10^{-5} mm.

Table A1. Model performance for the different ODE solvers/methods available in SciPy.

ODE solver	Method	Runtime (s)	ψ RMSE (m)	MB bias (mm)	MB RMSE (mm)
ode	BDF	1.71	0.000	-0.018	8.06×10^{-5}
ode	LSODA	1.96	0.000	0.001	7.11×10^{-5}
odeint	LSODA	2.69	0.000	0.000	7.04×10^{-5}
solve_ivp	BDF	7.45	0.001	-0.594	3.33×10^{-3}
solve_ivp	LSODA	2.44	0.000	0.059	6.12×10^{-4}

an erroneous approach, resulting in mass balance errors of more than 100 mm for our problem. For the solver flux output method of calculating the boundary fluxes, it is necessary to output states at reporting steps, and therefore it is not possible to use the discrete approximation for $C(\psi)$. The results in Table A3 all use calculation-step information. A more subtle issue is the order of the temporal integrator used by the ODE solver, which can be specified by the user. Here, we use either first-order or (variable) higher-order (as determined by the ODE solver) temporal integration methods. For the solver flux output method, we use higher-order temporal integration. The results are given in Table A3.

We see in Table 3 that the discrete $C(\psi)$ approach works quite well for first-order integration methods but is very slow. When higher-order integration methods are used, the model is faster, but the mass balance is chronically degraded. We think that this happens because with higher-order methods the model states evolve in a more complex manner (i.e., non-linear manner) over a calculation time step, so the linear approximation in Eq. (A1) is not good. It is noteworthy that the modeled ψ values were slightly modified using the discrete high-order approach. For comparison purposes, we looked at using analytical representations of $C(\psi)$ with first-order and higher-order methods, and this time the higher-order methods performed better. Overall, we recommend against using discrete $C(\psi)$ approximations, unless using a tailor-made ODE solver (such as Kavetski et al., 2001, 2002a, b).

A4 Alternative approaches to defining the Jacobian

As described in Sect. 4.1, providing the ODE solver with information about the Jacobian matrix is reported to improve the solution efficiency. Here we compare three approaches: no information provided about the Jacobian, defining the Jacobian pattern, and defining the full Jacobian matrix. For the last case, this was complex to define for our method, and therefore it was implemented for the high-order reporting-step solution procedure described in Sect. 4.3.2. The results are reported in Table A4.

We see that for both model configurations, defining the banded Jacobian sparsity pattern matrix led to improvements in performance of around 20%. This is modest, but because it is trivial to define the banded matrix, this is worthwhile. For the reporting-step model, when we defined the full Jacobian matrix, this led to a very slight improvement in performance over the banded solution (1.58 s vs. 1.63 s). Defining the full Jacobian is challenging and requires an additional function/function call in the code – we therefore recommend against using the full Jacobian matrix and recommend instead defining the banded matrix.

Table A2. Model performance using different calculation methods for the boundary fluxes.

Method		Runtime (s)	ψ RMSE (m)	MB bias (mm)	MB RMSE (mm)
SFOM		1.71	0.000	−0.018	8.06×10^{-5}
Reporting step	central	1.66	0.000	0.119	2.64×10^0
Reporting step	forward	1.66	0.000	0.119	5.29×10^0
Reporting step	backward	1.66	0.000	0.119	5.92×10^{-2}
Calculation step	central	5.24	0.000	0.046	1.13×10^{-4}
Calculation step	forward	5.24	0.000	0.046	3.24×10^{-3}
Calculation step	backward	5.24	0.000	0.046	3.14×10^{-3}

Table A3. Model performance using different approaches to calculate $C(\psi)$.

Method		Runtime (s)	ψ RMSE (m)	MB bias (mm)	MB RMSE (mm)
Analytical $C(\psi)$	solver flux output method	1.71	0.000	−0.018	8.06×10^{-5}
Analytical $C(\psi)$	first order	36.25	0.000	−0.49	4.97×10^{-4}
Analytical $C(\psi)$	high order	5.17	0.000	0.046	3.14×10^{-3}
Discrete $C(\psi)$	first order	44.62	0.000	0.317	1.92×10^{-7}
Discrete $C(\psi)$	high order	7.2	0.003	11.939	9.72×10^{-3}

Table A4. Model performance using different approaches to define the Jacobian matrix.

Method		Runtime (s)	ψ RMSE (m)	MB bias (mm)	MB RMSE (mm)
Solver flux output method	Jacobian pattern	1.71	0	−0.018	8.06×10^{-5}
	no Jacobian	2.17	0	−0.015	7.53×10^{-5}
Reporting-step flux calculation method	no Jacobian	2.16	0	0.119	5.92×10^{-2}
	Jacobian pattern	1.63	0	0.119	5.92×10^{-2}
	full Jacobian	1.58	0	0.123	5.92×10^{-2}

Table A5. Model performance with and without Numba JIT compilation.

Configuration	Runtime (s)	ψ RMSE (m)	MB bias (mm)	MB RMSE (mm)
With Numba	1.71	0.00	−0.018	8.06×10^{-5}
Without Numba	26.07	0.00	−0.018	8.06×10^{-5}

A5 Running the model with and without Numba JIT compilation

The best model configuration was also run with and without the Numba JIT compiler, and the result is shown in Table A5. It can be seen that Numba has no impact on the model output (accuracy and mass balance are identical for each run) as expected, but using Numba improves the runtime by a factor of ~ 15 . All other model runs reported in this paper use Numba.

Code and data availability. All of the scripts developed in this study are available from <https://github.com/amireson/openRE> (last access: 24 January 2023), release v1.0.1,

<https://doi.org/10.5281/zenodo.7497133> (Ireson, 2022). The code is written in Python and MATLAB and run using makefiles, which reproduce Figs. 2–7.

Author contributions. AMI conceived of this study, wrote all the scripts (except the pseudospectral similarity solution, written in MATLAB by SAM), performed the analysis, and drafted the manuscript. SAM came up with the idea for the proposed solver flux output method (SFOM). MPC, RJS, and SAM assisted with the study design and implementation and edited the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. Support for this study was provided by the Global Institute for Water Security and the Canada First Research Excellence Fund's Global Water Futures program.

Review statement. This paper was edited by Ludovic Räss and reviewed by James Craig and one anonymous referee.

References

- Bear, J. and Cheng, A. H. D.: Modeling groundwater flow and contaminant transport, Vol. 23, in: *Theory and Applications of Transport in Porous Media*, Springer, Dordrecht, 834, <https://doi.org/10.1007/978-1-4020-6682-5>, 2010.
- Beven, K., and Germann, P.: Macropores and water flow in soils revisited, *Water Resour. Res.*, 49, 3071–3092, <https://doi.org/10.1002/wrcr.20156>, 2013.
- Brown, P. N., Hindmarsh, A. C., and Byrne, G. D.: VODE. Variable Coefficient ODE Solver, *SIAM J. Sci. Stat. Comp.*, 10, 1038–1051, <https://doi.org/10.1137/0910062>, 1989.
- Celia, M. A., Bouloutas, E. T., and Zarba, R. L.: A general mass-conservative numerical solution for the unsaturated flow equation, *Water Resour. Res.*, 26, 1483–1496, <https://doi.org/10.1029/WR026i007p01483>, 1990.
- Clark, M. P., Fan, Y., Lawrence, D. M., Adam, J. C., Bolster, D., Gochis, D. J., Hooper, R. P., Kumar, M., Leung, L. R., Mackay, D. S., and Maxwell, R. M.: Improving the representation of hydrologic processes in Earth System Models, *Water Resour. Res.*, 51, 5929–5956, 2015.
- Clark, M. P., Zolfaghari, R., Green, K. R., Trim, S., Knoben, W. J. M., Bennett, A., Nijssen, B., Ireson, A., and Spiteri, R. J.: The Numerical Implementation of Land Models: Problem Formulation and Laugh Tests, *J. Hydrometeorol.*, 22, 1627–1648, <https://doi.org/10.1175/JHM-D-20-0175.1>, 2021.
- Farthing, M. W. and Ogden, F. L.: Numerical Solution of Richards' Equation: A Review of Advances and Challenges, *Soil Sci. Soc. Am. J.*, 81, 1257–1269, <https://doi.org/10.2136/sssaj2017.02.0058>, 2017.
- Goudarzi, S., Mathias, S. A., and Gluyas, J. G.: Simulation of three-component two-phase flow in porous media using method of lines, *Transport Porous Med.*, 112, 1–19, <https://doi.org/10.1007/s11242-016-0639-5>, 2016.
- Ireson, A. M.: openRE, v1.0.1, Zenodo [code], <https://doi.org/10.5281/zenodo.7497133>, 2022.
- Ireson, A. M. and Butler, A. P.: A critical assessment of simple recharge models: application to the UK Chalk, *Hydrol. Earth Syst. Sci.*, 17, 2083–2096, <https://doi.org/10.5194/hess-17-2083-2013>, 2013.
- Ireson, A. M., Mathias, S. A., Wheeler, H. S., Butler, A. P., and Finch, J.: A model for flow in the Chalk unsaturated zone incorporating progressive weathering. A model for flow in the Chalk unsaturated zone incorporating progressive weathering, *J. Hydrol.*, 365, 244–260, <https://doi.org/10.1016/j.jhydrol.2008.11.043>, 2009.
- Jackson, M. (Ed.): *Software Carpentry: Automation and Make*, Version 2016.06, <https://github.com/swcarpentry/make-novice> (last access: 24 January 2023), June 2016.
- Kavetski, D., Binning, P., and Sloan, S. W.: Adaptive time stepping and error control in a mass conservative numerical solution of the mixed form of Richards equation, *Adv. Water Resour.*, 24, 595–605, [https://doi.org/10.1016/S0309-1708\(00\)00076-2](https://doi.org/10.1016/S0309-1708(00)00076-2), 2001.
- Kavetski, D., Binning, P., and Sloan, S. W.: Adaptive backward Euler time stepping with truncation error control for numerical modelling of unsaturated fluid flow, *Int. J. Numer. Meth. Eng.*, 53, 1301–1322, <https://doi.org/10.1002/nme.329>, 2002a.
- Kavetski, D., Binning, P., and Sloan, S. W.: Noniterative time stepping schemes with adaptive truncation error control for the solution of Richards equation: NONITERATIVE TIME STEPPING SCHEMES, *Water Resour. Res.*, 38, 29–1–29–10, <https://doi.org/10.1029/2001WR000720>, 2002b.
- Lam, S. K., Pitrou, A., and Seibert, S.: Numba: A LLVM-based Python JIT compiler, *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, SC15: The International Conference for High Performance Computing, Networking, Storage and Analysis*, Austin, Texas, 15 November 2015, 1–6, <https://doi.org/10.1145/2833157.2833162>, 2015.
- Mathias, S. A. and Sander, G. C.: Pseudospectral methods provide fast and accurate solutions for the horizontal infiltration equation, *J. Hydrol.*, 598, 126407, <https://doi.org/10.1016/j.jhydrol.2021.126407>, 2021.
- Mathias, S. A., Skaggs, T. H., Quinn, S. A., Egan, S. N., Finch, L. E., and Oldham, C. D.: A soil moisture accounting-procedure with a Richards' equation-based soil texture-dependent parameterization, *Water Resour. Res.*, 51, 506–523, <https://doi.org/10.1002/2014WR016144>, 2015.
- Miller, C. T., Williams, G. A., Kelley, C. T., and Tocci, M. D.: Robust solution of Richards' equation for nonuniform porous media, *Water Resour. Res.*, 34(, 2599–2610, <https://doi.org/10.1029/98WR01673>, 1998.
- Milly, P. C. D.: A mass-conservative procedure for time-stepping in models of unsaturated flow, *Adv. Water Resour.*, 8, 32–36, [https://doi.org/10.1016/0309-1708\(85\)90078-8](https://doi.org/10.1016/0309-1708(85)90078-8), 1985.
- Milly, P. C. D.: A Mass-Conservative Procedure for Time-Stepping in Models of Unsaturated Flow, in: *Finite Elements in Water Resources*, edited by: Laible, J. P., Brebbia, C. A., Gray, W., and Pinder, G., Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-662-11744-6_9, pp. 103–112, 1984.
- Petzold, L.: Automatic Selection of Methods for Solving Stiff and Nonstiff Systems of Ordinary Differential Equations, *SIAM J. Sci. Stat. Comp.*, 4, 136–148, <https://doi.org/10.1137/0904010>, 1983.
- Rathfelder, K. and Abriola, L. M.: Mass conservative numerical solutions of the head-based Richards equation, *Water Resour. Res.*, 30, 2579–2586, <https://doi.org/10.1029/94WR01302>, 1994.
- Šimůnek, J., van Genuchten, M. Th., and Šejna, M.: The Hydrus-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. Version 3.0, HYDRUS Software Series 1, Department of Environmental Sciences, University of California Riverside, Riverside, CA, 2005.

- Šimůnek, J., van Genuchten, M. Th., and Šejna, M.: Recent developments and applications of the HYDRUS computer software packages, *Vadose Zone J.*, 15, 25, <https://doi.org/10.2136/vzj2016.04.0033>, 2016.
- Tocci, M. D., Kelley, C. T., and Miller, C. T.: Accurate and economical solution of the pressure-head form of Richards' equation by the method of lines, *Adv. Water Resour.*, 20, 1–14, [https://doi.org/10.1016/S0309-1708\(96\)00008-5](https://doi.org/10.1016/S0309-1708(96)00008-5), 1997.
- Tubini, N. and Rigon, R.: Implementing the Water, HEat and Transport model in GEOframe (WHETGEO-1D v.1.0): algorithms, informatics, design patterns, open science features, and 1D deployment, *Geosci. Model Dev.*, 15, 75–104, <https://doi.org/10.5194/gmd-15-75-2022>, 2022.
- van Genuchten, M.: A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils I, *Soil Sci. Soc. Am. J.*, 44, 892–898, <https://doi.org/10.2136/sssaj1980.03615995004400050002x>, 1980.
- Van Genuchten, M. T. H. and Gray, W. G.: Analysis of some dispersion corrected numerical schemes for solution of the transport equation, *Int. J. Numer. Meth. Eng.*, 12, 387–404, <https://doi.org/10.1002/nme.1620120302>, 1978.
- Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, T., Vanderborght, J., Young, M. H., Amelung, W., Aitkenhead, M., Allison, S. D., Assouline, S., Baveye, P., Berli, M., Brüggemann, N., Finke, P., Flury, M., Gaiser, T., Govers, G., Ghezzehei, T., Hallett, P., Hendricks Franssen, H. J., Heppell, J., Horn, R., Huisman, J. A., Jacques, D., Jonard, F., Kollet, S., Lafolie, F., Lamorski, K., Leitner, D., McBratney, A., Minasny, B., Montzka, C., Nowak, W., Pachepsky, Y., Padarian, J., Romano, N., Roth, K., Rothfuss, Y., Rowe, E. C., Schwen, A., Šimůnek, J., Tiktak, A., Van Dam, J., van der Zee, S. E. A. T. M., Vogel, H. J., Vrugt, J. A., Wöhling, T., Young, I. M.: Modeling Soil Processes: Review, Key Challenges, and New Perspectives, *Vadose Zone J.*, 15, [vzj2015.09.0131](https://doi.org/10.2136/vzj2015.09.0131), <https://doi.org/10.2136/vzj2015.09.0131>, 2016.