arXiv:2310.02806v1 [math.NA] 4 Oct 2023

A Data-facilitated Numerical Method for Richards Equation to Model Water Flow Dynamics in Soil

Zeyuan Song, Zheyu Jiang
School of Chemical Engineering
Oklahoma State University
Stillwater, OK 74074
{taekwon.song,zjiang}@okstate.edu

Abstract

Root-zone soil moisture monitoring is essential for precision agriculture, smart irrigation, and drought prevention. Modeling the spatiotemporal water flow dynamics in soil is typically achieved by solving a hydrological model, such as the Richards equation which is a highly nonlinear partial differential equation (PDE). In this paper, we present a novel data-facilitated numerical method for solving the mixed-form Richards equation. This numerical method, which we call the D-GRW (Data-facilitated global Random Walk) method, synergistically integrates adaptive linearization scheme, neural networks, and global random walk in a finite volume discretization framework to produce accurate numerical solutions of the Richards equation with guaranteed convergence under reasonable assumptions. Through three illustrative examples, we demonstrate and discuss the superior accuracy and mass conservation performance of our D-GRW method and compare it with benchmark numerical methods and commercial solver.

Keywords: Soil moisture, Richards equation, global random walk, neural network, finite volume method

1 Introduction

Root-zone soil moisture monitoring is essential for improving agricultural production and crop productivity, providing basis for precision irrigation and agriculture, preventing leaching of agro-chemicals and soil nutrients into groundwater, and predicting agricultural droughts [1]. Recent studies reveal that adjusting irrigation activities based on root-zone soil moisture information can reduce irrigation water consumption by 40-60% [2] and increase farmer's revenue by 20-60% [3]. Modeling the spatiotemporal behavior of root zone soil moisture from surface or near-surface soil moisture data is typically achieved by solving an agro-hydrological model that describes water movement through unsaturated soils. Nowadays, most agro-hydrological models are based on the Richards equation [4]:

$$\partial_t \Theta(\psi) + \nabla \cdot \mathbf{q} = 0, \tag{1a}$$

$$\mathbf{q} = -K(\Theta(\psi))\nabla(\psi + z), \tag{1b}$$

where ψ stands for pressure head (e.g., in m), $\mathbf{q}$ represents the water flux (e.g., in $\mathrm{m}^3/\mathrm{m}^2 \cdot \mathrm{s}$), Θ denotes the soil moisture content (e.g., in $\mathrm{m}^3/\mathrm{m}^3$), K is unsaturated hydraulic water conductivity (e.g., in m/s), $t \in [0, T]$ denotes the time (e.g., in s), and z corresponds to the vertical depth (e.g., in m). For unsaturated flow, both Θ and K are highly nonlinear functions of pressure head

ψ and soil properties, which makes Equations (1a) and (1b) very difficult to solve in practice. Specifically, $\Theta(\psi)$ and $K(\psi)$ are commonly referred to as the water retention curve and hydraulic conductivity function, respectively. These nonlinear relations have been regressed and tabulated as empirical models for various soil types. Some of the commonly used empirical models, such as the Haverkamp model [5], Mualem-van Genuchten model [6, 7], and Gardner's correlations [8], are given in Table 1.

Correlations	Hydraulic Conductivity Function	Water Retention Curve
Haverkamp [5]	$K(\psi) = K_s \frac{A}{A+\lvert\psi\rvert^{\gamma}}$	$\Theta(\psi) = \Theta_r + \frac{\alpha(\Theta_s-\Theta_r)}{\alpha+\lvert\psi\rvert^{\beta}}$
Mualem-van Genuchten [6, 7]	$K(\Theta) = K_s\sqrt{\frac{\Theta-\Theta_r}{\Theta_s-\Theta_r}}\cdot \left\{1-\left[1-(\frac{\Theta-\Theta_r}{\Theta_s-\Theta_r})^{\frac{n}{n-1}}\right]^{\frac{n-1}{n}}\right\}^2$	$\Theta(\psi) = \Theta_r + \frac{\Theta_s-\Theta_r}{\left[1+(\alpha\lvert\psi\rvert)^n\right]^{\frac{n-1}{n}}}$
Gardner [8]	$K(\psi) = K_s e^{\alpha\psi}$	$\Theta(\psi) = \Theta_r + (\Theta_s-\Theta_r)e^{\alpha\psi}$

Table 1: Empirical correlations of water retention curve and hydraulic conductivity function used in this work. In these models, A, γ, α, β, n, Θ_s, and Θ_r are soil-dependent parameters.

Since the analytical solutions of the Richards equation generally do not exist [9], most Richards equation solvers rely on discretizations to obtain numerical solutions. In particular, methods based on finite difference and finite element discretizations have been extensively studied and implemented [10, 11]. Nevertheless, they face challenges when handling large-scale problems and are prone to instability issues. Meanwhile, Ireson et al. [12] introduced a way to convert the 1-D Richards equation into an ordinary differential equation (ODE) using the method of lines. And the resulting ODE can be solved using the finite difference method. Apart from the aforementioned methods, the finite volume method (FVM) has demonstrated promising potential in achieving high solution accuracy and preserving the mass conservation when solving the Richards equation [13]. For example, Lai and Ogden [14] obtained a family of mass-conservative finite volume predictor-corrector solutions for the 1-D Richards equation. A second-order accurate monotone FVM was also proposed by Misiats and Lipnikov [15] for solving the 1-D Richards equation. Several other FVM-based numerical methods have also been developed to solve 1-D and 2-D Richards equations [16, 17, 18]. However, conventional FVM schemes typically convert the discretized Richards equation into a matrix equation. However, the resulting matrix is usually large, sparse and stiff, thereby posing computational challenges when solving the matrix equation.

To address this computational challenge, iterative methods such as the linearization scheme (L-scheme) have been adopted to solve the Richards equation [19]. For example, Bergamaschi and Putti [20] implemented a L-scheme under the mixed finite element method framework to solve the 2-D Richards equation. Standard linearization schemes typically adopt a fixed linearization parameter L for all time steps and discretized cells. Choosing an appropriate static value of L is not straightforward, as soil moisture content and pressure head vary with space and time. To addresss this limitation, Mitra and Pop [21] modified standard L-schemes for solving 1-D nonlinear diffusion equations by allowing L to be adaptive. More recently, Albuja and Avila [22] modified the Newton-type scheme and proposed a new linearization scheme for solving the 1-D Richards equation with guaranteed global convergence. Nevertheless, some of the existing adaptive L-schemes have been reported to be suffered from numerical oscillations [23], and so far, no adaptive L-scheme has been reported to successfully solve the 3-D Richards equations.

Meanwhile, the random walk (RW) model has been introduced as a promising approach to

improve the accuracy and stability of discretization-based numerical solvers. In the RW model, particles, in our case water molecules, can move to their neighboring positions following certain probabilities. Furthermore, random walks are memoryless, meaning that the current movements of particles are independent of their past movements. The stochastic and memoryless properties of random walks make the RW model an attractive tool for solving various transport problems [24], including the Richards equation [25] and its inverse problem [26]. However, in the conventional RW model, only one particle is allowed to move at a time, and particles cannot stay at the same location. To overcome this drawback, the global random walk (GRW) model was recently proposed [27, 28], which allows all particles to move and/or stay at the same time. The GRW model has been successfully implemented to model groundwater transport [27, 29], diffusion process [30, 29], and to solve the 1-D and 2-D Richards equation [28]. Some of these GRW model implementations are further built upon the L-scheme framework [28]. Nevertheless, they still adopt a static constant linearization parameter L. In addition, a key assumption used in existing GRW implementations for solving the Richards equation is that the pressure head is proportional to the number of particles in a discretized cell, which is yet to be verified. Last but not least, successful extension of the GRW model to solving the 3-D Richards equation has not been reported in the literature.

Recently, we proposed a data-facilitated global random walk (D-GRW) framework combining the GRW model and neural networks to successfully solve the 1-D Richards equation [31]. Specifically, the use of neural networks results in more accurate characterization of the relationship between the number of water molecules and the pressure head. Later, we extended this work [31] to solve the 3-D Richards equations [32]. In this work, we extend our prior works and introduce several innovative solutions to improve the numerical accuracy and computational efficiency of our D-GRW method, including:

- We generalize our prior works [31, 32] to solve any d-dimensional ($d = 1, 2, 3$) Richards equation and provide theoretical justifications on existence, uniqueness, and convergence behavior of solutions.
- We introduce a novel coarse-to-fine approach to enhance the solution accuracy of our D-GRW method without requiring a large amount of high-accuracy, fine-mesh training data. We demonstrate that this coarse-to-fine approach maintains a good balance between computational efficiency and solution accuracy.
- Through systematic case studies, we show that, by synergistically integrating adaptive L-scheme, global random walk, and neural network, our D-GRW method significantly enhances the performance of FVM-based solver in preserving the underlying physics and mass conservation associated with the Richards equation.

We organize the subsequent sections of the paper as follows. In Section 2, we discuss the well-posedness of continuous and discretized Richards equation. Next, in Section 3, we derive the FVM-based adaptive L-scheme formulation and discuss its convergence behavior. Then, in Section 4, we incorporate the adaptive L-scheme formulation in the D-GRW method. To compare the performance of our D-GRW method with benchmark solvers, we present illustrative examples in Section 5. Finally, we summarize our key findings and discuss future directions in Section 6.

2 Well-posedness of the Richards Equation

The performance of a numerical PDE solver is theoretically governed by the well-posedness of the original PDE [33]. A problem is said to be well-posed if its weak solution exists, is unique, and

depends continuously on the problem's initial conditions [33, 34]. As a result, well-posedness is an essential property that certifies the accuracy and reliability of numerical solutions to the Richards equation. Misiats and Lipnikov [15] studied the existence and uniqueness of the weak solution of a simplified Richards equation. Later, Abdellatif et al. [35] showed the well-posedness of continuous and discretized Richards equation. In this section, we will summarize prior works on the well-posedness of the continuous and discretized Richards equation [15, 21, 35, 36] and highlight key results.

2.1 Well-posedness of the Continuous Richards Equation

We consider a general d-dimensional field $\Omega \subset \mathbb{R}^d$ ($d = 1, 2, 3$) with finite measure. The Richards equation (Equations (1a) and (1b)) with homogeneous Dirichlet boundary condition and initial condition can be written as:

$$\begin{cases} \partial_t \Theta(\psi) - \nabla \cdot (K(\Theta)\nabla(\psi + z)) = 0, \\ \psi(\cdot, \mathbf{x}) = 0 \text{ on } \partial\Omega, \\ \psi(0, \cdot) = \psi_0(\cdot) \in H_0^1(\Omega) \text{ over } \Omega \times [0, T]. \end{cases} \tag{2}$$

Following Misiats and Lipnikov [15], the weak solution of Equation (2) is defined as follows:

Definition 2.1 (modified from Definition 1 of [15])**.** *Given $\psi \in L^2\left([0,T], H_0^1(\Omega)\right)$, if for any $v \in L^2\left([0,T], H_0^1(\Omega)\right)$ such that $\partial_t v \in L^\infty\left([0,T], H_0^1(\Omega)\right)$ and $v(T, \cdot) = 0$,*

$$\int_0^T \int_\Omega \left(\Theta(\psi) - \Theta(\psi_0)\right) \partial_t v - K(\Theta)\nabla(\psi + z)\nabla \cdot v \, \mathrm{d}\Omega \, \mathrm{d}t = 0$$

holds, then ψ is the weak solution of Equation (2)*.*

In Definition 2.1, note that $\partial_t v \in L^\infty\left([0,T], H_0^1(\Omega)\right)$ can be embedded into any $L^p\left([0,T], H_0^1(\Omega)\right)$ for $1 \leq p < \infty$ using Hölder's inequality. To see the well-posedness of the continuous Richards equation, we first adopt the Kirchhoff's transformation to Equation (2) [13, 37]. The Kirchhoff transformation $\mathcal{K}$ performs the following mapping:

$$\mathcal{K} : L^2\left([0,T], H_0^1(\Omega)\right) \to L^2\left([0,T], H_0^1(\Omega)\right), \tag{3a}$$

$$\psi \mapsto \int_0^\psi K\left(\Theta(\omega)\right) \mathrm{d}\omega. \tag{3b}$$

To simplify the notations, let $u \triangleq \int_0^\psi K\left(\Theta(\omega)\right) \mathrm{d}\omega$, $\mathcal{M}(u) \triangleq \Theta(\mathcal{K}^{-1}(u))$, and $\mathcal{L}\left(\nabla u, \mathcal{M}(u)\right) \triangleq \nabla u + \mathcal{K}\left(\mathcal{M}(u)\right)\nabla z$. Note that $\nabla u = K\left(\Theta(\psi)\right)\nabla\psi$ from the chain rule. With this, Equation (2) can be rewritten as:

$$\begin{cases} \partial_t \mathcal{M}(u) - \nabla \cdot \mathcal{L}\left(\nabla u, \mathcal{M}(u)\right) = 0, \\ u(\cdot, \mathbf{x}) = 0 \quad \text{on } \partial\Omega, \\ u(0, \cdot) = u_0(\cdot) \triangleq \int_0^{\psi_0} K\left(\Theta(\omega)\right) \mathrm{d}\omega, \end{cases} \tag{4}$$

where ψ_0 is the same initial condition as in Equation (2). We now introduce two results from [38].

Lemma 2.1 (modified from the main theorem of [38])**.** *Consider an elliptic-parabolic equation:*

$$\partial_t \mathcal{M}(u) - \nabla \cdot \mathcal{L}\left(\nabla u, \mathcal{M}(u)\right) = 0, \tag{5}$$

with initial condition:

$$\mathcal{M}\left(u(0, \cdot)\right) = \mathcal{M}\left(u_0(\cdot)\right), \tag{6}$$

and boundary condition:

$$u(\cdot, \mathbf{x}) = 0 \quad \text{on } \partial\Omega, \tag{7}$$

then there exists a unique weak solution u given the initial condition of Equation (6).

We remark that, since water retention curve models are monotonically non-decreasing functions of pressure head, $\mathcal{M}$ is monotone and continuous. Thus, Equation (6) reduces to $u(0, \cdot) = u_0(\cdot)$. Furthermore, after applying the Kirchholf transformation, the Richards equation has the same form as Equation (5).

Lemma 2.2 (Lemma 2.2 of [38])**.** *Let u be a weak solution of Equation* (5) *with initial condition of Equation* (6) *and boundary condition of Equation* (7)*, then*

$$\operatorname*{ess\,lim}_{t\to 0} \int_\Omega (\mathcal{M}(u) - \mathcal{M}(u_0))^+ \, \mathrm{d}\Omega = 0,$$
$$\operatorname*{ess\,lim}_{t\to 0} \int_\Omega (\mathcal{M}(u_0) - \mathcal{M}(u))^+ \, \mathrm{d}\Omega = 0.$$

In other words, Lemma 2.2 implies that $\lim_{t\to 0} \int_\Omega (\mathcal{M}(u) - \mathcal{M}(u_0)) \, \mathrm{d}\Omega = 0$ holds for the weak solution u almost everywhere on the domain Ω. In Theorem 2.3 below, we state the well-posedness of the continuous Richards equation, which can be derived from Lemmas 2.1 and 2.2 [15, 38].

Theorem 2.3. *Equations* (2) *and* (4) *are well-posed. That is, the weak solutions $\psi(\cdot, \cdot)$ and $u(\cdot, \cdot)$ are unique and only depend on the initial conditions $\psi_0(\cdot)$ and $u_0(\cdot)$, respectively.*

To understand Theorem 2.3, recall from Lemma 2.1 that the unique weak solution $u(\cdot, \cdot)$ of Equation (4) exists given initial condition $u_0(\cdot)$. Meanwhile, the dependence of the solution $u(\cdot, \cdot)$ on initial condition $u_0(\cdot)$ is presented in Lemma 2.2. This implies that the weak solution of Equation (4) is well-posed. Next, to show that Equation (2) is also well-posed, we first remark that, given two weak solutions u_1 and u_2 of Equation (4), the corresponding weak solutions of Equation (2) are $\psi_1 = \mathcal{K}^{-1}(u_1)$ and $\psi_2 = \mathcal{K}^{-1}(u_2)$, respectively. Furthermore, one can verify that:

$$K_l|\psi_1 - \psi_2| = \Big| \int_{\psi_2}^{\psi_1} K_l \, \mathrm{d}\omega \Big| \leq \Big| \int_{\psi_2}^{\psi_1} K\left(\Theta(\omega)\right) \mathrm{d}\omega \Big| = \big|\mathcal{K}(\psi_1) - \mathcal{K}(\psi_2)\big| = |u_1 - u_2|, \tag{8}$$

where $K_l = \min_{\psi \in [\psi_2, \psi_1]} K(\Theta(\psi))$. Thus, for any $\varepsilon > 0$, there exists $\delta = K_l \varepsilon$ such that when $|u_1 - u_2| < \delta$, $|\psi_1 - \psi_2| < \varepsilon$, which implies the uniqueness of weak solution of Equation (2). Lastly, to derive the continuous dependence of weak solution ψ on its initial condition ψ_0 for Equation (2), we start from the continuous dependence property of Equation (4), which suggests that small deviations in the initial conditions would result in small deviations in the weak solution. Thus, for any $\varepsilon' > 0$, there exists $\delta' = M\varepsilon'$ $(0 < M < \infty)$ such that $|u - \hat{u}| < \varepsilon'$ given that $|u_0 - \hat{u}_0| < \delta'$. Here, u_0 and $\hat{u}_0$ are the initial conditions of Equation (4) associated with the weak solutions u and $\hat{u}$, respectively. From Equation (8), one can directly extend the continuous dependence property of Equation (4) to Equation (2) by letting $\delta'' = \frac{\delta'}{K_l}$ and $\varepsilon'' = \frac{\varepsilon'}{K_l}$ and showing that for any $\varepsilon'' > 0$, there exists δ'' such that $|\psi - \hat{\psi}| < \varepsilon''$ when $|\psi_0 - \hat{\psi}_0| < \delta''$. At this point, we obtain the well-posedness of Equation (2).

2.2 Well-posedness of the Discretized Richards Equation

Now, we examine the discretized version of Equation (2) obtained by spatial discretization into N small cells as well as temporal discretization using implicit Euler method on the time domain with

time step size Δt. We denote $\psi(m\Delta t, \cdot)|_{V_i}$ as $\psi_i^m(\cdot)$ for every time step $m = 0, 1, \ldots, \lceil \frac{T}{\Delta t} \rceil - 1$ and cell $i = 1, \ldots, N$. Then, the discretized version of Equation (2) is formulated as:

$$\begin{cases} \Theta(\psi_i^{m+1}) - \Theta(\psi_i^m) - \Delta t \nabla \cdot \left[K\left(\Theta(\psi_i^{m+1})\right) \nabla \left(\psi_i^{m+1} + z\right) \right] = 0, \\ \psi_j(\cdot) = 0 \quad \text{for } V_j \subset \partial\Omega, \\ \psi(0, \cdot) = \psi_0(\cdot). \end{cases} \tag{9}$$

Next, we define an inner product $\langle \cdot, \cdot \rangle_{V_i} : L^2\left([0,T], H_0^1(\Omega)\right) \to L^2\left([0,T], H_0^1(\Omega)\right)$ as $\langle f, g \rangle_{V_i} \triangleq \int_{V_i} fg \,\mathrm{d}V$. When V_i is sufficiently small, $\int_{V_i} fg \,\mathrm{d}V = (fg)_i \mathrm{vol}(V_i)$. We use $(fg)_i$ to denote the value of fg evaluated at cell i, and $\|f\|^2 \triangleq \langle f, f \rangle_{V_i} = (f^2)_i \mathrm{vol}(V_i)$. Then, the weak solution of the discretized version of Equation (9) is defined as follows:

Definition 2.2. *Given $\psi_i^m \in H_0^1(\Omega)$, if for any $v \in H_0^1(\Omega)$ and $v(T, \cdot) = 0$,*

$$\left\langle \left(\Theta(\psi^{m+1}) - \Theta(\psi^m)\right), v \right\rangle_{V_i} + \Delta t \left\langle K\left(\Theta(\psi^{m+1})\right) \nabla(\psi^{m+1} + z), \nabla v \right\rangle_{V_i} = 0 \tag{10}$$

holds, then ψ_i^{m+1} is a weak solution of the discretized Richards equation of Equation (9).

To see how Equation (10) is originated from the discretized Richards equation of Equation (9), we observe that:

$$\begin{aligned} \langle \nabla \cdot [K\left(\Theta(\psi^{m+1})\right) \nabla \left(\psi^{m+1} + z\right)], v \rangle_{V_i} &= \int_{V_i} \nabla \cdot \left[K\left(\Theta(\psi^{m+1})\right) \nabla \left(\psi^{m+1} + z\right) \right] v \,\mathrm{d}V \\ &= \int_{\partial V_i} K\left(\Theta(\psi^{m+1})\right) \nabla \left(\psi^{m+1} + z\right) v \,\mathrm{d}S - \int_{V_i} K\left(\Theta(\psi^{m+1})\right) \nabla \left(\psi^{m+1} + z\right) \nabla v \,\mathrm{d}V \\ &= -\langle K\left(\Theta(\psi^{m+1})\right) \nabla(\psi^{m+1} + z), \nabla v \rangle_{V_i}, \end{aligned}$$

in which we use integration by parts for higher dimensions and the surface integral over ∂V_i is 0.

Now, similar to the continuous version of the Richards equation, we can apply the Kirchhoff transformation $\mathcal{K}$ to Equation (10) and obtain:

$$\left\langle \mathcal{M}(u^{m+1}) - \mathcal{M}(u^m), v \right\rangle_{V_i} + \Delta t \langle \nabla u^{m+1}, \nabla v \rangle_{V_i} + \Delta t \left\langle \mathcal{K}\left(\mathcal{M}(u^{m+1})\right) \nabla z, \nabla v \right\rangle_{V_i} = 0, \tag{11}$$

where $u_i^m \triangleq \int_0^{\psi_i^m} K\left(\Theta(\omega)\right) \mathrm{d}\omega$ and $\mathcal{M}(u_i^m) \triangleq \Theta(\mathcal{K}^{-1}(u_i^m))$ for discretized cell i and time step m.

The well-posedness of weak solution $u_i^{m+1} \in H_0^1(\Omega)$ of Equation (11) has been shown in [21, 36] given $u_i^m \in H_0^1(\Omega)$. Specifically, given $u_i^m \in H_0^1(\Omega)$, there exists a sequence $\{u_i^{m+1,s}\}_s$ for $s = 1, 2, \ldots$ that converges to the unique weak solution u_i^{m+1}. Therefore, there correspondingly exists a sequence $\{\psi_i^{m+1,s}\}_s \in H_0^1(\Omega)$ for $s = 1, 2, \ldots$ that converges to a unique $\psi_i^{m+1} = \mathcal{K}^{-1}(u_i^{m+1})$. Such a sequence can be obtained by the iterative numerical scheme (e.g., the adaptive L-scheme introduced in Section 3), whose convergence is indicated by:

$$|\psi_i^{m+1,s} - \psi_i^{m+1}| < \epsilon \qquad \forall s \geq S, \tag{12}$$

in which the iteration number S depends on the tolerance parameter ϵ. In other words, given the converged solution ψ_i^m for time step m, one can move to the next time step $m+1$ and obtain a unique weak solution ψ_i^{m+1} of Equation (9) as the limit of sequence $\{\psi_i^{m+1,s}\}_s$. Thus, the discretized Richards equation of Equation (9) is well-posed. Next, in Section 3, we will study the convergence behavior of $\{\psi_i^{m+1,s}\}_s$ under the adaptive L-scheme formulation of the discretized Richards equation.

3 Finite Volume Discretization Using Adaptive Linearization Procedure

In this section, we introduce the adaptive L-scheme formulation of the Richards equation discretized by FVM. This results in the sequence $\{\psi_i^{m,s}\}_s$. We will analyze the convergence behavior of $\{\psi_i^{m,s}\}_s$.

3.1 Adaptive L-scheme of Richards Equation

We present the key steps involved in the discretization of the Richards equation using FVM, followed by applying adaptive L-scheme to solve the discretized equation iteratively. Let us consider a control volume $V \subset \Omega$. Integrating both sides of Equation (1a) over V gives:

$$\int_V \partial_t \Theta(\psi)\,\mathrm{d}V = \int_V \nabla \cdot \big[K(\Theta)\nabla(\psi + z)\big]\,\mathrm{d}V. \tag{13}$$

We can apply the divergence theorem to convert the volume integral on the RHS of Equation (13) into a surface integral:

$$\partial_t \Theta(\psi)\mathrm{vol}(V) = \oint_{S_V} K(\Theta)\nabla(\psi + z) \cdot \mathbf{n}\,\mathrm{d}S_V, \tag{14}$$

where $\mathrm{vol}(V)$ is the volume of V, S_V is the surface of V and $\mathbf{n}$ is the outward pointing unit normal to the boundary ∂V. The control volume V is discretized into N small cells $i = 1, \ldots, N$. We denote the common surface shared by cell i and cell j as $\omega_{i,j}$. We denote the operator $K(\cdot)\nabla(\cdot)$ and the outward pointing unit normal vector $\mathbf{n}$ on $\omega_{i,j}$ as $\big[K(\cdot)\nabla(\cdot)\big]_{\omega_{i,j}}$ and $\mathbf{n}_{\omega_{i,j}}$, respectively. With this, the spatial discretization of Equation (14) is expressed as:

$$\partial_t \Theta_i \mathrm{vol}(V_i) = \sum_{j \in \mathcal{N}_i} \big[K(\Theta)\nabla(\psi + z)\big]_{\omega_{i,j}} \cdot \mathbf{n}_{\omega_{i,j}} A_{\omega_{i,j}} \qquad \forall i = 1, \ldots, N, \tag{15}$$

where $\partial_t \Theta_i$ refers to the time derivative $\partial_t \Theta(\psi_i)$ on the cell i, $\mathcal{N}_i$ denotes the index set of all the neighboring cells sharing a common surface with i, and $A_{\omega_{i,j}}$ is the area of surface $\omega_{i,j}$.

In standard L-scheme, for every cell i and time step $m+1$, one would add the term $L(\psi_i^{m+1,s+1} - \psi_i^{m+1,s})$ to either side of Equation (15) to solve the Richards equation in an iterative manner. We remark that not all choices of L would lead to the same solution quality. In fact, some choices of L yield more accurate solutions than others. In standard L-schemes, which adopt a static parameter L, a trial-and-error procedure is typically required to obtain the "optimal" single L value. Not only is this search procedure tedious to implement, the solutions obtained are also less accurate most of the time as we will show in Section 5.1. To overcome the limitation of standard L-scheme, we adopt an adaptive L-scheme inspired by Mitra and Pop [21] by adding the term $L_i^{m+1,s}(\psi_i^{m+1,s+1} - \psi_i^{m+1,s})$ to the LHS of Equation (15) and rearranging it as:

$$\psi_i^{m+1,s+1} = \psi_i^{m+1,s} + \frac{1}{L_i^{m+1,s}} \sum_{j \in \mathcal{N}_i} \big[K(\Theta)\nabla(\psi + z)\big]_{\omega_{i,j}}^{m+1,s} \cdot \mathbf{n}_{\omega_{i,j}} A_{\omega_{i,j}} - \frac{1}{L_i^{m+1,s}} \partial_t \Theta_i^{m+1} \mathrm{vol}(V_i), \tag{16}$$

where the linearization parameter $L_i^{m+1,s}$ depends on the specific discretized cell, time step, and iteration count. Here, the time derivative $\partial_t \Theta_i^{m+1}$ is discretized as:

$$\partial_t \Theta_i^{m+1} = \frac{\Theta(\psi_i^{m+1,s}) - \Theta(\psi_i^m)}{\Delta t}, \tag{17}$$

where ψ_i^m is the converged water pressure head for cell i at time step m. Combining Equations (16) and (17) leads to the adaptive L-scheme formulation of the Richards equation discretized by FVM:

$$\begin{aligned}\psi_i^{m+1,s+1} = \psi_i^{m+1,s} &+ \frac{1}{L_i^{m+1,s}} \sum_{j\in\mathcal{N}_i} K_{\omega_{i,j}}^{m+1,s} \frac{(\psi+z)_j^{m+1,s} - (\psi+z)_i^{m+1,s}}{\text{dist}(V_j, V_i)} \mathbf{e}\cdot\mathbf{n}_{\omega_{i,j}} A_{\omega_{i,j}} \\ &- \frac{1}{L_i^{m+1,s}} \frac{\Theta(\psi_i^{m+1,s}) - \Theta(\psi_i^m)}{\Delta t} \text{vol}(V_i),\end{aligned} \tag{18}$$

where $\mathbf{e} = (1,1,1)$, and $\text{dist}(\cdot,\cdot)$ represents the Euclidean distance function.

We would like to point out that, in practice, $L_i^{m+1,s}$ needs to be sufficiently large because otherwise, the convergence of adaptive L-scheme could become slow, as the RHS of Equation (18) would approach infinity. In actual implementation, to prevent $L_i^{m+1,s}$ from being too close to 0 in case when $\dot{\Theta}(\psi_i^{m+1,s})$ approaches 0, we impose a lower bound on $L_i^{m+1,s}$:

$$L_i^{m+1,s} \geq L_l \quad \text{if} \max_{\widetilde{s}\leq s} \dot{\Theta}(\psi_i^{m+1,\widetilde{s}}) \leq \rho, \tag{19}$$

where $\rho > 0$ is a user-specified parameter [21].

3.2 Convergence Analysis of Adaptive L-scheme

Here, we analyze the the convergence behavior of our adaptive L-scheme for the Richards equation. From Bergamaschi and Putti [20], as $L_i^{m+1,s}$ approaches $\dot{\Theta}(\psi_i^{m+1,s}) \triangleq \frac{\mathrm{d}\Theta}{\mathrm{d}\psi}|_{\psi_i^{m+1,s}}$, our adaptive L-scheme essentially becomes the Newton's scheme which exhibits quadratic convergence. In general, the convergence of our adaptive L-scheme formulation of Equation (18) is characterized in Theorem 3.1. To show this, the idea is to find $\psi_i^{m+1,s+1} \in H_0^1(\Omega)$ given ψ_i^m and $\psi_i^{m+1,s} \in H_0^1(\Omega)$ such that:

$$\begin{aligned}\left\langle \Theta(\psi^{m+1,s+1}) - \Theta(\psi^m), v\right\rangle_{V_i} &+ \Delta t L_i^{m+1,s}(\psi_i^{m+1,s+1} - \psi_i^{m+1,s})v \\ &= -\Delta t \Big\langle K\left(\Theta(\psi^{m+1})\right) \nabla(\psi^{m+1,s+1} + z), \nabla v\Big\rangle_{V_i}\end{aligned} \tag{20}$$

holds for any $v \in H_0^1(\Omega)$.

Theorem 3.1. *$\{\psi_i^{m+1,s}\}_s$ converges to $\psi_i^{m+1} \in H_0^1(\Omega)$ for $m = 0, 1, \ldots, \lceil \frac{T}{\Delta t} \rceil - 1$ and $i = 1, \ldots, N$.*

Proof. First, we state two general observations for water infiltration in soil:

O-1. There exists a scaling factor $0 < \gamma_{s,0} < \infty$ such that $||\psi_i^{m+1,s+1} - \psi_i^{m+1}|| < \gamma_{s,0}||\psi_i^{m+1,s} - \psi_i^{m+1}||$. In other words, $||\psi_i^{m+1,s+1} - \psi_i^{m+1}|| < +\infty$.

O-2. $\dot{\Theta}(\psi)$ is positive which is valid in commonly used water retention curve models.

First, we subtract Equation (10) from Equation (20) and obtain:

$$\begin{aligned}\left\langle \Theta(\psi^{m+1,s+1}) - \Theta(\psi^{m+1}), v\right\rangle_{V_i} &+ \Delta t L_i^{m+1,s}(\psi_i^{m+1,s+1} - \psi_i^{m+1,s})v \\ &= -\Delta t \Big\langle K\left(\Theta(\psi^{m+1})\right) \nabla(\psi^{m+1,s+1} - \psi^{m+1}), \nabla v\Big\rangle_{V_i}.\end{aligned} \tag{21}$$

Since $\psi_i^{m+1,s+1} - \psi_i^{m+1,s} = (\psi_i^{m+1,s+1} - \psi_i^{m+1}) - (\psi_i^{m+1,s} - \psi_i^{m+1})$, Equation (21) can be rewritten as:

$$\begin{aligned}\Delta t L_i^{m+1,s}(\psi_i^{m+1,s} - \psi_i^{m+1})v = &\left\langle \Theta(\psi^{m+1,s+1}) - \Theta(\psi^{m+1}), v\right\rangle_{V_i} + \Delta t L_i^{m+1,s}(\psi_i^{m+1,s+1} - \psi_i^{m+1})v \\ &+ \Delta t\left\langle K\left(\Theta(\psi^{m+1})\right)\nabla(\psi^{m+1,s+1} - \psi^{m+1}), \nabla v\right\rangle_{V_i}.\end{aligned} \tag{22}$$

Let $v = \psi_i^{m+1,s+1} - \psi_i^{m+1}$, then from O-1, the LHS of Equation (22) can be bounded by:

$$\begin{aligned}\Delta t L_i^{m+1,s}(\psi_i^{m+1,s} - \psi_i^{m+1})v &= \Delta t L_i^{m+1,s}(\psi_i^{m+1,s} - \psi_i^{m+1})(\psi_i^{m+1,s+1} - \psi_i^{m+1}) \\ &< \frac{\gamma_{s,0}\Delta t L_i^{m+1,s}}{\text{vol}(V_i)}(\psi_i^{m+1,s} - \psi_i^{m+1})^2\text{vol}(V_i) \\ &= \frac{\gamma_{s,0}\Delta t L_i^{m+1,s}}{\text{vol}(V_i)}||\psi_i^{m+1,s} - \psi_i^{m+1}||^2.\end{aligned} \tag{23}$$

Similarly, for the second term on the RHS of Equation (22), we have:

$$\begin{aligned}\Delta t L_i^{m+1,s}(\psi_i^{m+1,s+1} - \psi_i^{m+1})v &= \Delta t L_i^{m+1,s}(\psi_i^{m+1,s+1} - \psi_i^{m+1})^2 \\ &= \frac{\Delta t L_i^{m+1,s}}{\text{vol}(V_i)}||\psi_i^{m+1,s+1} - \psi_i^{m+1}||^2.\end{aligned} \tag{24}$$

By the mean value theorem, the first term on the RHS of Equation (22) can be written as:

$$\begin{aligned}\left\langle \Theta(\psi^{m+1,s+1}) - \Theta(\psi^{m+1}), v\right\rangle_{V_i} &= \dot{\Theta}(\xi)\left\langle \psi^{m+1,s+1} - \psi^{m+1}, v\right\rangle_{V_i} \\ &= \dot{\Theta}(\xi)||\psi_i^{m+1,s+1} - \psi_i^{m+1}||^2,\end{aligned} \tag{25}$$

for $\xi \in (\psi_i^{m+1,s+1}, \psi_i^{m+1})$.

Lastly, for the third term on the RHS of Equation (22), we have:

$$\Delta t\left\langle K\left(\Theta(\psi^{m+1})\right)\nabla(\psi^{m+1,s+1} - \psi^{m+1}), \nabla v\right\rangle_{V_i} = \Delta t K\left(\Theta(\psi_i^{m+1})\right)||\nabla\psi_i^{m+1,s+1} - \nabla\psi_i^{m+1}||^2. \tag{26}$$

Combining Equations (23) through (26) leads to:

$$\begin{aligned}||\psi_i^{m+1,s+1} - \psi_i^{m+1}|| &< \sqrt{\frac{\gamma_{s,0}L_i^{m+1,s}\Delta t}{L_i^{m+1,s}\Delta t + \dot{\Theta}(\xi)\text{vol}(V_i)}}||\psi_i^{m+1,s} - \psi_i^{m+1}|| \\ &< \sqrt{\gamma_{s,0}}||\psi_i^{m+1,s} - \psi_i^{m+1}||.\end{aligned} \tag{27}$$

Let $\gamma_{s,1} \triangleq \sqrt{\gamma_{s,0}}$ and replace O-1 by $\gamma_{s,1}$. One can repeat the derivations above to obtain $||\psi_i^{m+1,s+1} - \psi_i^{m+1}|| < \sqrt{\gamma_{s,1}}||\psi_i^{m+1,s} - \psi_i^{m+1}||$. We can further define $\gamma_{s,2} = \sqrt{\gamma_{s,1}}$, and so on, until we obtain a scaling factor $\gamma_s < 1$ such that $||\psi_i^{m+1,s+1} - \psi_i^{m+1}|| < \gamma_s||\psi_i^{m+1,s} - \psi_i^{m+1}||$.

With this, one can show that for a given tolerance $\epsilon > 0$, there exists $S \in \mathbb{N}^+$ such that:

$$\begin{aligned}||\psi_i^{m+1,s+1} - \psi_i^{m+1}|| &< \gamma_s||\psi_i^{m+1,s} - \psi_i^{m+1}|| < \gamma_s\gamma_{s-1}||\psi_i^{m+1,s-1} - \psi_i^{m+1}|| < \cdots \\ &< \prod_{k=1}^{s}\gamma_k||\psi_i^{m+1,1} - \psi_i^{m+1}|| < \left(\max_{1\le k\le s}\gamma_k\right)^s||\psi_i^{m+1,1} - \psi_i^{m+1}|| < \epsilon \quad \forall s \ge S,\end{aligned}$$

which is consistent with the convergence criterion of Equation (12). This completes the proof. □

4 Data-facilitated Global Random Walk (D-GRW) Framework

In this section, we develop our novel D-GRW method to enhance the solution accuracy and mass conservation performance of existing GRW-based solvers. We discuss how the D-GRW method is implemented to solve the adaptive L-scheme formulation of the Richards equation discretized by FVM. Later, we will illustrate the attractiveness of D-GRW method through case studies.

4.1 Formulation and Convergence of D-GRW Method

To formulate the D-GRW method for the discretized Richards equation, let $n_i^{m,s}$ be the number of water molecules in cell i at iteration step s and time step m, and $\delta n_{i,j}^{m,s}$ be the number of particles moving from the cell i to the neighboring cell $j \in \mathcal{N}_i$. Clearly, we have:

$$n_i^{m,s} = \delta n_{i,i}^{m,s} + \sum_{j \in \mathcal{N}_i} \delta n_{j,i}^{m,s}. \tag{28}$$

For instance, for the 1-D case, we can write $n_i^{m,s} = \delta n_{i,i}^{m,s} + \delta n_{i+1,i}^{m,s} + \delta n_{i-1,i}^{m,s}$ for every $i = 1, \ldots, N$. Thus, as long as the relationship between pressure head $\psi_i^{m,s}$ and the number of particles $n_i^{m,s}$ is established, Equation (28) can be incorporated into our adaptive L-scheme formulation of Equation (18) to solve the Richards equation iteratively. As pointed out earlier, a key assumption made in existing GRW methods [28] is that $\psi_i^{m,s}$ is proportional to $n_i^{m,s}$ for any iteration and time step. Although this assumption is shown to be valid for diffusion equations [30], we will show that, for the Richards equation, the actual relationship may not be smooth or explicit.

Having recognized this, we propose a data-driven approach in D-GRW method to accurately approximate the nonlinear mapping and inverse mapping between $\psi_i^{m,s}$ and $n_i^{m,s}$ using two multi-layer perceptrons (MLPs). MLP is capable of approximating any function provided that the neural network contains enough neurons [39, 40]. Depending on the specific problem to be solved, the desired choices of optimal optimizer, number of hidden layers, and activation functions can vary. Based on our extensive search and hyperparameter tuning, we find that a simple three-layer MLP with 256 neurons in each layer achieves the best performance for most 1-D through 3-D problems compared to other more complex neural network architectures (e.g., LSTM). Also, we find that stochastic gradient decent (SGD) optimizer often outperforms others (e.g., Adam or RMSProp). The mapping and inverse mapping obtained from the MLPs are given by $f : n_i^{m,s} \to \psi_i^{m,s}$ and $f^{-1} : \psi_i^{m,s} \to n_i^{m,s}$, respectively. Substituting the trained inverse mapping f^{-1} to Equation (18) gives the D-GRW framework for the adaptive L-scheme formulation:

$$n_i^{m+1,s+1} = n_i^{m+1,s} + \frac{1}{L_i^{m+1,s}} \sum_{j \in \mathcal{N}_i} K_{\omega_{i,j}}^{m+1,s} \mathbf{e} \cdot \mathbf{n}_{\omega_{i,j}} \frac{n_j^{m+1,s} - n_i^{m+1,s}}{\text{dist}(V_j, V_i)} A_{\omega_{i,j}} + f^{-1}(J), \tag{29}$$

where $J = \frac{1}{L_i^{m+1,s}} \sum_{j \in \mathcal{N}_i} K_{\omega_{i,j}}^{m+1,s} \mathbf{e} \cdot \mathbf{n}_{\omega_{i,j}} \frac{z_j^{m+1,s} - z_i^{m+1,s}}{\text{dist}(V_j, V_i)} A_{\omega_{i,j}} - \frac{1}{L_i^{m+1,s}} \frac{\Theta_i^{m+1,s} - \Theta_i^m}{\Delta t} \text{vol}(V_i)$. With this, one can extend Theorem 3.1 and derive a similar convergence property for $\{n_i^{m+1,s}\}_s$:

Theorem 4.1. *$\{n_i^{m+1,s}\}_s$ converges to $n_i^{m+1} \in H_0^1(\Omega)$ for $m = 0, 1, \ldots, \lceil \frac{T}{\Delta t} \rceil - 1$ and $i = 1, \ldots, N$.*

4.2 Implementation of D-GRW Method

Two parts, namely neural network training and solution process, are involved in implementing our proposed D-GRW method to solve the Richards equation. In neural network training, two MLPs are

trained to learn the mapping and inverse mapping between $\psi_i^{m,s}$ and $n_i^{m,s}$, that is, f and f^{-1}, respectively. First, for each cell i and time step m, we obtain $n_i^{m,S}$ from GRW-based solvers (e.g., [28]) and the corresponding $\psi_i^{m,S}$ from conventional non-adaptive L-scheme formulation, where S is a pre-specified total iteration number. The resulting solution pairs, $\left(\psi_i^{m,S}, n_i^{m,S}\right)$, form the original reference solutions. Then, to every reference solution, we will add Gaussian noise $Z_q \sim \mathcal{N}(0, \sigma_q^2)$ with different variances $\sigma_1^2, \ldots, \sigma_Q^2$. In other words, $\psi_{i,q}^{m,S} \triangleq \psi_i^{m,S} + Z_q$ and $n_{i,q}^{m,S} \triangleq n_i^{m,S} + Z_q$. After this data augmentation step, the resulting expanded set of reference solutions $\left\{(\psi_{i,q}^{m,S}, n_{i,q}^{m,S})\right\}_{i,m,q}$ will be used for MLP training. In Section 5.1, we will show that introducing Gaussian noise is an effective data augmentation technique that reduces the biases of reference solutions and enhances generalization performance [41], thereby significantly improving the accuracy of numerical solutions. Once offline training (only one time) is complete, the initial and boundary conditions of the Richards equation will be mapped to the number of particles using the trained MLP for the inverse mapping f^{-1}. Then, starting from $m = 0$ (initial condition), Equation (29) is solved iteratively with the help of trained f^{-1} to obtain a sequence $\{n_i^{m+1,s}\}_s$ for each new time step $m+1$ and cell i. To monitor the convergence of $\{n_i^{m+1,s}\}_s$, we define the relative error RE as:

$$\mathrm{RE} \triangleq \max_s \left\{ \left| \frac{n_i^{m+1,s+1} - n_i^{m+1,s}}{n_i^{m+1,s+1}} \right| \right\}. \tag{30}$$

Once RE is reduced below the set tolerance tol, we declare convergence of $\{n_i^{m+1,s}\}_s$ to n_i^{m+1}. With this, we can determine the converged ψ_i^{m+1} from the MLP approximating mapping f, followed by obtaining other physical quantities such as soil moisture level Θ_i^{m+1} and $\mathbf{q}_i^{m+1}$ from Table 1 and Equation (1b).

5 Case Studies

In this section, we systematically validate our D-GRW numerical framework on selected 1-D through 3-D benchmark problems modified from the literature [11, 42, 43]. In the 1-D case study adopted from Celia et al. [11], we compare the solution accuracy of our D-GRW method with other benchmark solvers. In the 2-D case study adopted from Gąsiorowski and Kolerski [42], we show that our novel numerical solver can better capture underlying physics and conservation laws of the Richards equation. Lastly, in the 3-D case study adopted from Tracy [43] in which an analytical solution to the Richards equation exists, we show that our numerical solutions have much lower relative error compared to the results obtained from existing GRW-based solvers.

5.1 One-dimensional Benchmark Problem

Here, we study the 1-D benchmark problem over a 40 cm deep soil presented by Celia et al. [11]. The hydraulic conductivity function and water retention curve follow Haverkamp et al. [5]'s model (see Table 1) whose parameters are listed in Table 5.1. And the initial condition and two boundary conditions of this case study are $\psi(z, 0) = -61.5$ cm and $\psi(40 \text{ cm}, t) = -20.7$ cm, $\psi(0, t) = -61.5$ cm, respectively [5].

To implement our D-GRW algorithm, both MLPs contain 3 hidden layers and 256 neurons in each layer. We use the leaky ReLU activation function [44] and stochastic gradient descent (SGD) optimizer with a learning rate of 0.001. To generate the reference solutions for training, we divide the spatial domain into 40 discretized cells (i.e., grid size $\Delta z = 1$ cm) using 41 mesh points and

Soil-dependent Parameters	Values	Units
Saturated hydraulic conductivity, K_s	0.00944	cm/s
Saturated soil moisture content, Θ_s	0.287	–
Residual soil moisture content, Θ_r	0.075	–
α in Haverkamp's model	1.611×10^6	cm
A in Haverkamp's model	1.175×10^6	cm
β in Haverkamp's model	3.96	–
γ in Haverkamp's model	4.74	–
Total time, T	360	s

Table 2: Soil property parameters and their values used in the 1-D case study of Celia et al. [11] based on the empirical model developed by Haverkamp et al. [5].

solve the Richards equation using both finite difference method of Celia et al. [11] and the GRW-based solver of Suciu et al. [28]. Specifically, we obtain 1640 pressure head solution data (41 mesh points $\times$ 40 time steps) following the method of Celia et al. [11]. Meanwhile, solution data (i.e., the number of particles) from the GRW solver [28] are obtained for different static linearization parameter (L) values ranging from 0.5 to 15. Next, we make multiple copies of the pressure head solution data, and for each L value, we append one copy to the GRW solutions. Then, we perform data augmentation to these appended reference solutions by adding Gaussian noises whose standard derivations vary from 0.1 to 0.5. Overall, this leads to a total of 17104 data points for MLP training and validation. Note that these data points are based on a coarse grid ($\Delta z = 1$ cm). Thus, they can be obtained relatively efficiently. On the other hand, in the solution step, we consider a more refined grid containing 101 mesh points ($\Delta z = 0.4$ cm). We consider two stopping scenarios. In Scenario 1, we set the error tolerance $\mathrm{tol} = 4 \times 10^{-5}$, whereas in Scenario 2, we set the total number of iterations $S = 20000$. We compare the accuracy of D-GRW method with other benchmark solvers against the "ground truth" solutions of Celia et al. [11] synthesized in a very dense grid under both scenarios.

Through this 1-D illustrative example, we will highlight the benefits of (a) adopting an adaptive L-scheme as opposed to standard L-scheme, (b) implementing the D-GRW method as opposed to GRW algorithms, and (c) integrating the adaptive L-scheme with D-GRW in a holistic numerical framework.

5.1.1 The Benefits of Adaptive L-scheme

To illustrate how adaptive L-scheme improves solution accuracy, we compare the pressure head profiles obtained at $t = T = 360$ sec for (a) standard L-scheme with static L values chosen from 1 to 10 and (b) our proposed adaptive L-scheme formulation of Equation (18). As shown in Figure 1, when using standard L-scheme, the choice of linearization parameter L can greatly impact the solution accuracy. In the absence of ground truth solutions, identifying the optimal L value for standard L-scheme will be challenging, which limits its use in solving new problems. On the other hand, our adaptive L-scheme of Equation (18) gives pressure head solutions that closely match with the ground truth solutions in one shot. Also, we point out that, by imposing the additional constraint on the lower bound of $L_i^{m+1,s}$ (see Equation (19)), our adaptive L-scheme successfully bypasses the singularity issue as $L_i^{m+1,s}$ approaches to 0 and correctly calculates the pressure head solutions for $z \in [0, 20 \text{ cm}]$ where $\dot{\Theta}$ becomes small.

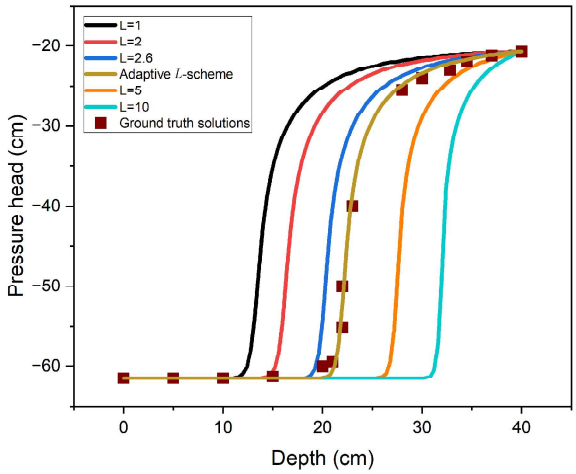

Figure 1: Comparison of pressure head profiles after 20000 iterations for the 1-D benchmark problem [11] using standard and adaptive L-schemes (Equation (18)). The ground truth solutions are obtained based on very fine space and time steps referred to as the "correct solutions" by Celia et al. [11] (only a selected data points are shown for figure clarity).

5.1.2 The Need for Data-facilitated Global Random Walk

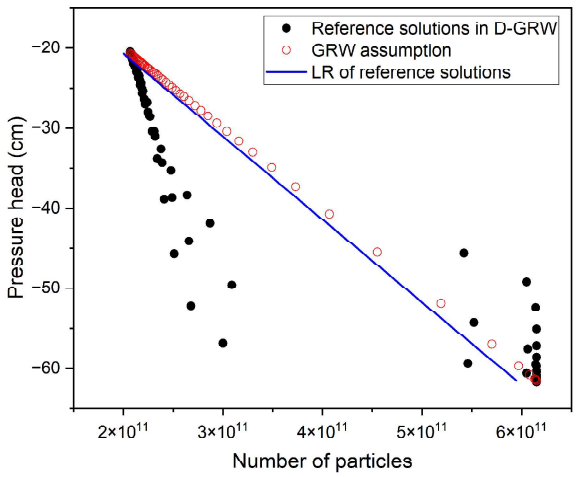

Figure 2: The relationship between $\psi_i^{m,S}$ and the number of particles $|n_i^{m,S}|$ at $S = 20000$ of reference solutions used in our D-GRW method (black filled circles), the GRW solver (red empty circles), and the linear regression of reference solutions using ordinary least square (blue line). The linear regression of reference solutions gives $|n_i^{m,s}| = -9.667 \times 10^9 \psi_i^{m,s}$, whereas GRW solver assumes $|n_i^{m,s}| = -1 \times 10^{10} \psi_i^{m,s}$.

As mentioned earlier, a key underlying assumption of existing GRW models is that $\psi_i^{m,s}$ is proportional to $n_i^{m,s}$. To validate this assumption, in Figure 2, we plot 1640 original reference solutions prior to performing data augmentation. To obtain the initial and boundary conditions in terms of the number of particles to be used by the GRW-based solution framework [28], we multiply these conditions, originally represented in terms of pressure head, by a factor of 10^{10} particles per cm of pressure head. It is clear from Figure 2 that, while the magnitude of pressure head shows roughly an increasing trend with respect to the number of particles present in a cell, the exact relationship between the number of particles and pressure head is neither smooth nor explicit. Interestingly, we do observe close-to-linear trends in specific regions of pressure head values (i.e., between -20 and -30 cm, and between -50 and -60 cm). Comparing with Figure 1, one can see that these two regions correspond to the two ends of the spatial domain. In fact, in both ends, $\frac{\partial \psi}{\partial z}$ is close to 0, indicating that the advection term in Equation (1b), $\nabla \cdot (K \nabla z) = \frac{\partial K}{\partial z} = \frac{\partial K}{\partial \psi} \frac{\partial \psi}{\partial z}$, vanishes. Thus, the Richards equation essentially becomes the diffusion equation, in which the proportionality assumption between pressure head and the number of particles is shown to be valid [30]. In other regions where pressure head changes rapidly with respect to depth, this assumption no longer holds, and one needs to seek other methods, such as neural network approximation used in our D-GRW method, to effectively learn the mapping and inverse mapping between pressure head and the number of particles.

5.1.3 Improving D-GRW Method Performance via Data Augmentation

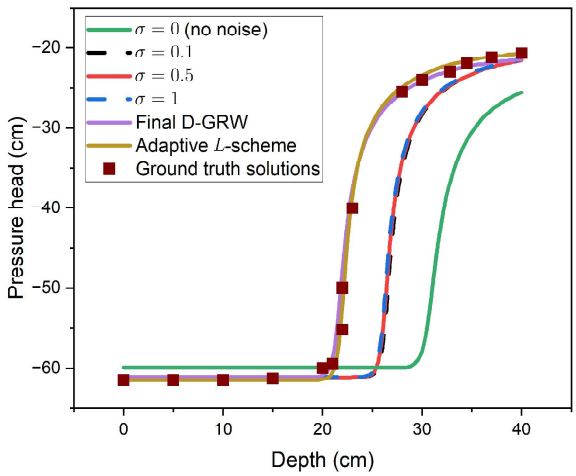

Figure 3: Pressure head profiles synthesized using different numerical methods and different training data. Specifically, curves named $\sigma = 0, 0.1, 0.5, 1$ are calculated from our D-GRW method, in which only 1640 reference solutions (with Gaussian noise added for $\sigma = 0.1, 0.5, 1$) are used for MLP training. The purple curve is the final D-GRW pressure head profile obtained by considering the full 17104 data points for MLP training through data augmentation. The brown curve only implements the adaptive L-scheme formulation of Equation (18) without incorporating the GRW model.

To understand the impact of data augmentation to the solution quality of our D-GRW method, we examine several scenarios shown in Figure 3. First, we observe that, compared to directly using

the original reference solutions for MLP training, simply introducing Gaussian noise to the reference solutions while keeping the size of training data the same can significantly improve the solution accuracy of our D-GRW method. This observation illustrates the need for introducing Gaussian noise to reference solutions for more accurate learning of MLPs. Second, we notice that that there is almost no difference nor further improvement in final pressure head profile when Gaussian noises of different magnitudes are directly added to the original reference solutions without performing data augmentation. Third, increasing the size of training data (from 1640 to 17104) via data augmentation of original reference solutions is an effective and efficient way to improve the solution accuracy of our D-GRW method, as the pressure head profile matches very well with the ground truth solution. Last but not least, we show the need for combining adaptive L-scheme formulation and data-facilitated global random walk model in our D-GRW method, since the pressure head profile obtained from our integrated numerical framework better agrees with ground truth solutions (especially between $z \in [25, 35]$ cm) compared to that obtained from directly solving Equation (18).

5.1.4 Comparison with Benchmark Solvers

To compare the computational efficiency of our D-GRW solver and with GRW-based algorithms, we calculate the relative error RE defined in Equation (30) under two stopping scenarios described before. Specifically, we examine three different approaches to obtain f and f^{-1} for solving Equation (29) (see Figure 2). As shown in Tables 3 and 4, while our proposed neural network approach requires slightly more iterations to reach $\mathrm{tol} = 4 \times 10^{-5}$, its convergence is accelerated as more iterations are performed, and at 20000 iterations, our D-GRW algorithm achieves the lowest relative error. By comparing the pressure head profiles in Figure 4a and b, it is clear that as iteration proceeds, the relative error of all iterative methods will decrease and solution accuracy will improve. In particular, after 20000 iterations, our D-GRW framework produces the pressure head profile that best matches with ground truth solutions.

f^{-1} in Equation (29) obtained by	Iterations	tol
Proportionality assumption (GRW)	$16,598$	3.9995×10^{-5}
Linear regression	$18,818$	3.9998×10^{-5}
MLP training (D-GRW)	$19,905$	3.9988×10^{-5}

Table 3: The number of iterations needed to reach the prespecified relative error of 4×10^{-5} under Scenario 1 for different global random walk implementations to solve the adaptive L-scheme of Equation (29).

f^{-1} in Equation (29) obtained by	Iterations	RE
Proportionality assumption (GRW)	$20,000$	3.8697×10^{-5}
Linear regression	$20,000$	3.9437×10^{-5}
MLP training (D-GRW)	$20,000$	3.5976×10^{-5}

Table 4: Relative error RE results of different global random walk implementations after 20000 iterations to solve the adaptive L-scheme of Equation (29) under Scenario 2.

In our earlier work [31], we illustrated the D-GRW method's ability to accurately capture underlying physics of water flow dynamics in soil, including the relationships among pressure head,

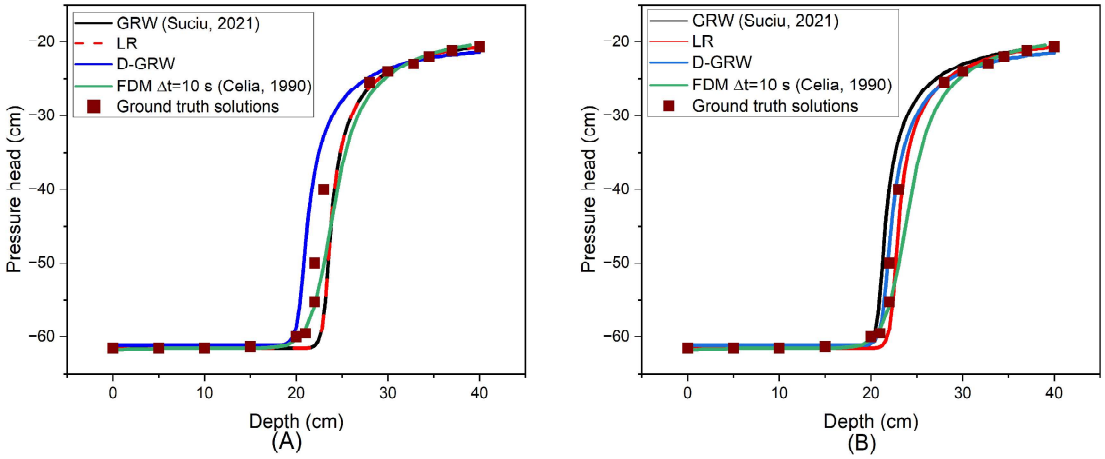

Figure 4: Pressure head profiles at $t = T = 360$ sec obtained by different algorithms under (a) Scenario 1, and (b) Scenario 2. Note that the finite difference method (FDM) by Celia et al. [11] is not an iterative method. As iteration proceeds, these pressure head profiles will move toward ground truth solutions.

soil moisture, and water flux. Extending these observations, we compare the mass conservation performance among various algorithms. We adopt the mass balance measure MB defined in [11]:

$$\mathrm{MB} = \frac{\text{total additional mass in the domain}}{\text{total water flux into the domain}}. \tag{31}$$

As summarized in Table 5, among all numerical methods studied, in both Scenarios 1 and 2, our D-GRW method achieves the highest MB values while using the coarsest time step, which is more computationally efficient than using finer time steps. Similarly, Table 6 shows that, compared to other numerical methods, our our D-GRW method achieves the highest MB values when fixing Δt to be the same. In addition, given more iterations, the MB value in all iterative methods will improve. After 20000 iterations (Scenario 2), the MB value reaches over 99.5% for all iterative methods. This result is also consistent with how the accuracy of the pressure head profiles improves with increasing iterations when comparing Figure 4a with 4b.

Method used	Scenario	Δt (sec)	MB
Linear regression	1	16.88	93.86%
GRW model [28]	1	16.86	93.98%
D-GRW method	1	21.54	**95.76%**
Linear regression	2	16.86	99.78%
GRW model	2	16.86	99.83%
D-GRW method	2	19.32	**99.91%**
Finite difference method [11]	N/A	10	95.00%

Table 5: MB results of different numerical methods implemented in 1-D case study. Here, Δt is determined for each method by the heuristic formula of Equation 5 in [28].

Method used	Scenario	MB ($\Delta t = 10$ sec)
Linear regression	1	95.49%
GRW model [28]	1	94.81%
D-GRW method	1	**96.04%**
Linear regression	2	99.80%
GRW model	2	99.85%
D-GRW method	2	**99.91%**
Finite difference method [11]	N/A	95.00%

Table 6: MB results of different numerical methods implemented in 1-D case study. Here, Δt is set to be 10 sec.

5.2 Two-dimensional benchmark problem

In the second example, we study the 2-D Richards equation for an infiltration process in a $1\text{m} \times 1\text{m}$ loam soil field [42]. The spatial steps in both horizontal (Δx) and vertical (Δz) directions are set to be 0.02 m. And the time step used for this comparison study is $\Delta t = 10$ seconds. The Mualem-van Genuchten model (see Table 1) was used in this case study. And the soil property parameters, given by Carsel and Parrish [45], are shown in Table 7 below:

Property	Symbol	Value	Units
Saturated hydraulic conductivity	K_s	2.89×10^{-6}	m/s
Saturated water content	Θ_s	0.43	–
Residual water content	Θ_r	0.078	–
van Genuchten Constant	α	3.6	m^{-1}
van Genuchten Constant	n	1.56	–
Total time	T	1.26×10^4	sec

Table 7: Soil-dependent parameters and constants used in 2-D case study.

The initial and boundary conditions of this case study are given by:

Initial condition: $\psi(x, z, t = 0 \text{ s}) = \begin{cases} 0 \text{ m}, & x \in [0.46, 0.54] \text{ m}, \ z = 0 \text{ m}, \\ -10 \text{ m}, & \text{otherwise.} \end{cases}$

Boundary condition: $\psi(x \in [0.46, 0.54] \text{ m}, z = 0 \text{ m}, t) = 0$ m, no slip conditions for other boundaries.

Note that the initial and boundary conditions are symmetric along $x = 0.5$ m. We first obtain 1734 original reference solutions for MLP training from the 2-D GRW solver [28], which implements the standard L-scheme ($L = 0.5$) and a spatial step of 0.05 m. Then, we apply data augmentation by adding Gaussian noises with σ^2 values ranging from 0.1 to 0.5 to generate a total of 26010 reference solutions (which also contain the original reference solutions). These reference solutions are used to train the two MLPs for our D-GRW method. Each MLP contains 2 hidden layers and 25 neurons in each layer. ReLU activation function is adopted in each layer, and each MLP is trained by Levenberg-Marquardt optimization for 1000 epochs. We set the total iteration number to be $S = 500$, at which the relative error calculated using Equation (30) for our D-GRW method and the GRW solver are given by 3.675×10^{-6} and 1.094×10^{-5}, respectively. This indicates that our D-GRW method achieves faster convergence per iteration than the GRW-based solver.

Meanwhile, we also simulate this 2-D problem using HYDRUS 2D software (version 5.0) [46] and compare the pressure head results at $t = T = 1.26 \times 10^4$ sec with our D-GRW method and the GRW solver. From Figure 5, we can draw two observations. First, the pressure head solution profiles for both GRW and D-GRW methods appear to be symmetric along $x = 0.5$ m, whereas HYDRUS 2D shows a clear asymmetric profile. As pointed out earlier, since the initial and boundary conditions are symmetric along $x = 0.5$ m, symmetry in the pressure head solutions is expected. This suggests that both GRW and D-GRW based solvers can capture some degree of underlying physics of the original problem. Second, despite the assymetric behavior in pressure head profile, the size of isolines for the HYDRUS 2D simulation result is more similar to our D-GRW solution than to the GRW solver solution. This observation is also consistent with the information presented in Figure 7a. In fact, both observations can be carried over to the soil moisture profile as well, as shown in Figures 6 and 7b.

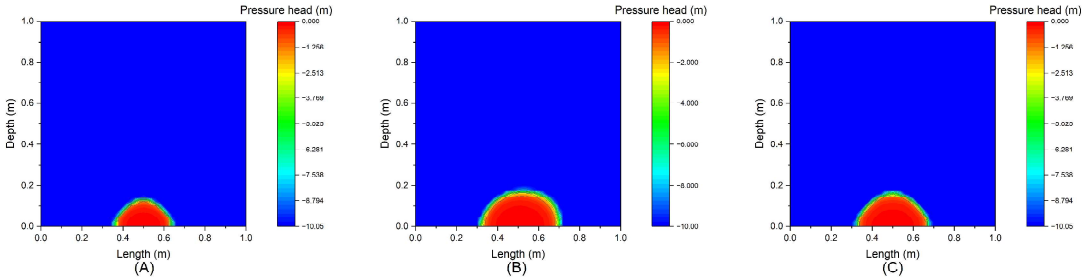

Figure 5: Pressure head solution profile obtained from three numerical methods: (a) GRW-based solver ($L = 0.5$), (b) HYDRUS 2D software, and (c) our D-GRW method.

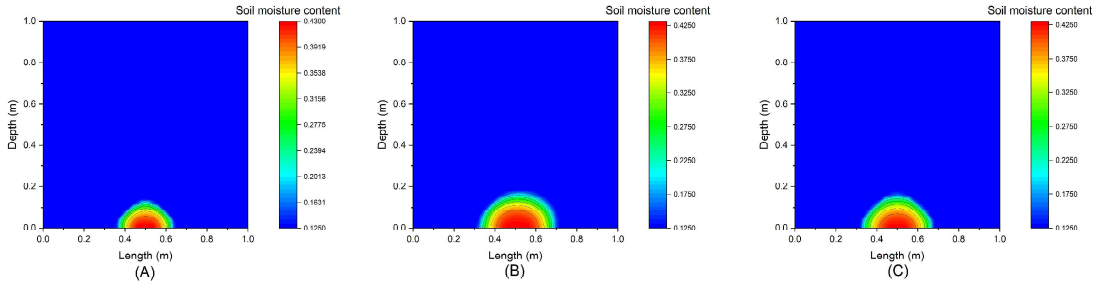

Figure 6: Soil moisture solution profile obtained from three numerical methods: (a) GRW-based solver ($L = 0.5$), (b) HYDRUS 2D software, and (c) our D-GRW method.

On the other hand, when comparing the water flux results, we see from Figure 8 that the GRW-based solution no longer preserves the symmetry of water flux profile along the horizontal direction. In other words, among the three numerical methods considered in this case study, our D-GRW method achieves the best performance in terms of preserving the symmetry implied by the problem. This result is also consistent with the mass conservation calculations using Equation (31), as our D-GRW method achieves significantly higher MB value compared to other benchmark solver (see Table 8).

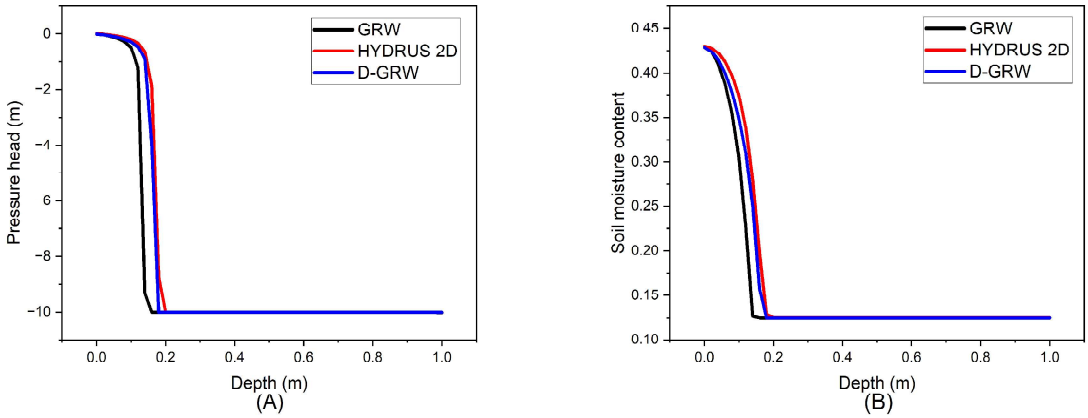

Figure 7: Cross-sectional view ($x = 0.5$m) of: (a) the pressure head profile, and (b) soil moisture profile.

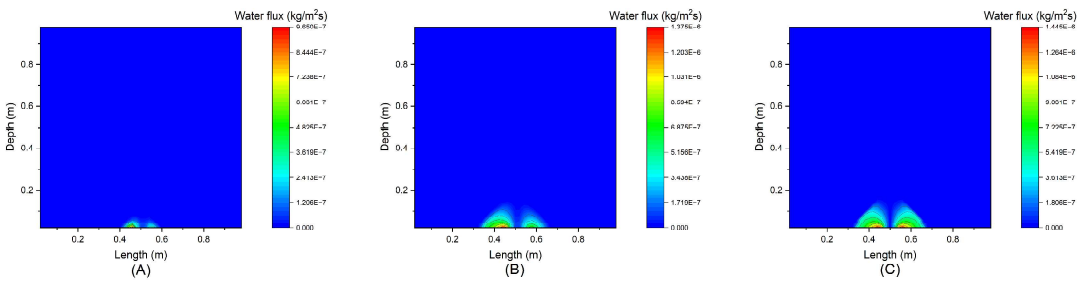

Figure 8: Magnitude of water flux along the horizontal (x-axis) direction for three numerical solvers: (a) GRW-based solver ($L = 0.5$), (b) HYDRUS 2D software, and (c) our D-GRW method. Note that, along the horizontal direction, the water flux is negative in $[0\text{ m}, 0.5\text{ m}]$ and positive in $(0.5\text{ m}, 1\text{ m}]$.

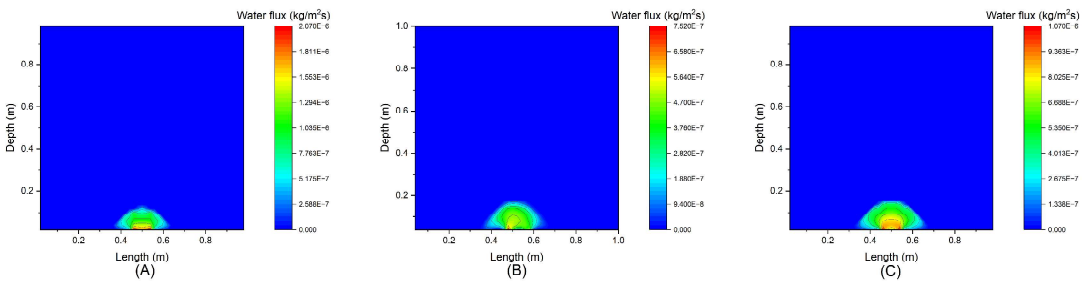

Figure 9: Magnitude of water flux along the vertical (z-axis) direction for three numerical solvers: (a) GRW-based solver ($L = 0.5$), (b) HYDRUS 2D software, and (c) our D-GRW method.

Method	MB ($\Delta t = 10$ sec)
GRW solver	63.93%
HYDRUS 2D simulation	62.45%
D-GRW method	**73.07%**

Table 8: MB values of three methods at $x = 0.5$m.

5.3 Three-dimensional benchmark problem

Lastly, we consider a 3-D water infiltration example where the analytical solution exists [43]. In this example, Ω is a 3-D cuboid $[0, a] \times [0, b] \times [0, c]$. The hydraulic conductivity function follows the Gardner's model [8] (see Table 1). The initial condition is given by:

$$\psi(x, y, z, t = 0) = h_r,$$

where h_r is a constant. And the boundary condition is given by:

$$\psi(x, y, z = c, t) = \frac{1}{\alpha} \ln \left[\exp(\alpha h_r) + \overline{h}_0 \sin \frac{\pi x}{a} \sin \frac{\pi y}{b} \right],$$

where $\overline{h}_0 = 1 - \exp(\alpha h_r)$. The pressure head solution for this problem was derived in [43] as:

$$\psi = \frac{1}{\alpha} \ln \left\{ \exp(\alpha h_r) + \overline{h}_0 \sin \frac{\pi x}{a} \sin \frac{\pi y}{b} \exp\left(\frac{\alpha(c-z)}{2} \right) \left[\frac{\sinh \beta z}{\sinh \beta c} + \frac{2}{zd} \sum_{k=1}^{\infty} (-1)^k \frac{\lambda_k}{\gamma} \sin(\lambda_k z) \exp(-rt) \right] \right\}, \tag{33}$$

where $d = \frac{\alpha(\Theta_s - \Theta_r)}{K_s}$, $\lambda_k = \frac{k\pi}{c}$, $\gamma = \frac{\lambda_k^2 + \beta^2}{c}$ and $\beta = \sqrt{\frac{\alpha^2}{4} + (\frac{\pi}{a})^2 + (\frac{\pi}{b})^2}$.

The infinite series in Equation (33) is convergent by alternating series test, and we consider only the first 1000 terms of this series. Note from Equation (33) that the analytical solution depends only on the saturated (Θ_s) and residual soil moisture content (Θ_s). And the Mualem-van Genuchten correlation [6, 7] tabulated in Table 1 was used for the water retention curve $\Theta(\psi)$. The constants and parameters used in this case study are listed in Table 9:

Property	Symbol	Value	Units
Saturated hydraulic conductivity	K_s	1.1	m/s
Saturated soil moisture	Θ_s	0.5	–
Residual soil moisture	Θ_r	0	–
Parameter in Gardner's model	α	0.1	m^{-1}
Parameter in intial and boundary conditions	h_r	-15.24	m
Length of Ω	a	2	m
Width of Ω	b	2	m
Depth of Ω	c	2	m
Total time	T	86400	sec

Table 9: Soil-dependent parameters and constants used in the 3-D case study.

Our goal is to compare the accuracy of our D-GRW method with GRW solvers using this analytical solution as the benchmark. To the best of our knowledge, there exists no GRW solver for 3-D Richards equation in the literature. Thus, we develop our own 3-D GRW-based solver in house

to allow for comparison. We obtain 1734 original reference solutions using the in-house developed 3-D GRW solver [28] that implements the standard L-scheme ($L = 0.5$). Then, data augmentation is applied by introducing Gaussian noise, resulting in a total of 8820 data points (which include the original reference solutions) for MLP training. For both GRW and D-GRW methods, we set the tolerance to be 10^{-9}, which can be achieved in less than 500 iterations for each time step.

We examine and compare the pressure head solutions at $z = 0.5$ and 1m, which are shown in Figure 10 and 11, respectively. We quantify the differences between the numerical solution and the analytical solution by $\frac{\psi_{\text{analytical}} - \psi_{\text{numerical}}}{\psi_{\text{analytical}}}$. From the relative difference heat map of Figure 10c,e and 11c,e, we observe that, first, the magnitude of relative error of our D-GRW based solver is significantly lower than that of GRW based solver. And second, the largest relative error of our D-GRW pressure head solution occurs around the four corners of the x–y domain, whereas the largest relative error of conventional GRW-based solution occurs in the center of the x–y domain. Furthermore, in each cell, the relative error of GRW-based pressure head solution is always non-positive, whereas that of D-GRW based pressure head solution can be positive or negative.

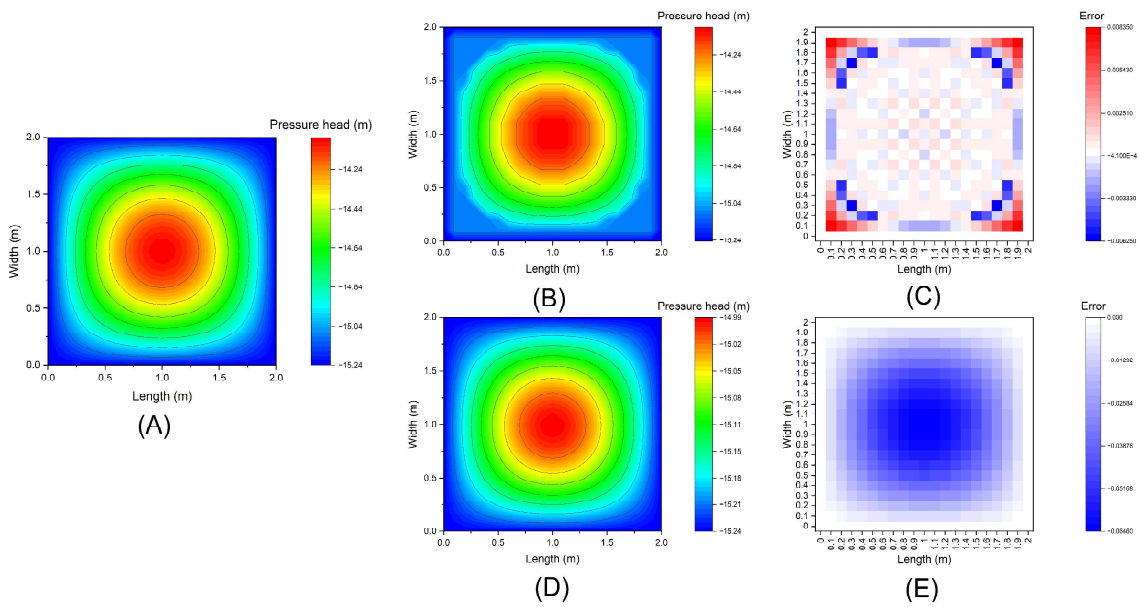

Figure 10: Pressure head solution at $z = 0.5$ m of different methods: (A) analytical solution, (B) D-GRW method, (C) the relative error between analytical and D-GRW based solutions, (D) GRW solver ($L = 0.5$) and (E) the relative error between analytical solution and GRW based solutions.

Here, we provide some justifications to these observations. First, for conventional GRW solver that embeds the standard L-scheme formulation, we observe from Equations (18) that:

$$\psi_{\text{analytical}} - \psi_{\text{numerical}} \propto \left\{ \sum_{j \in \mathcal{N}_i} \left[K(\psi) \nabla (\psi + z) \right]_{\omega_{i,j}}^{m+1,s} \cdot \mathbf{n}_{\omega_{i,j}} A_{\omega_{i,j}} - \partial_t \Theta_i^{m+1} \text{vol}(V_i) \right\},$$

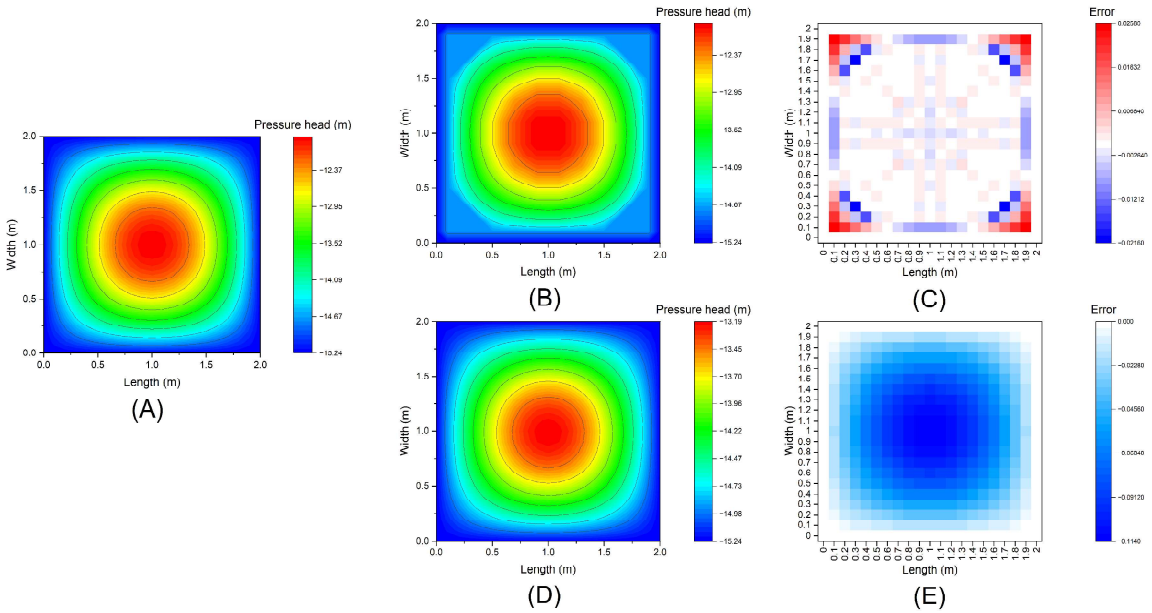

Figure 11: Pressure head solution at $z = 1.0$ m of different methods: (A) analytical solution, (B) D-GRW method, (C) the relative error between analytical and D-GRW based solutions, (D) GRW solver ($L = 0.5$) and (E) the relative error between analytical solution and GRW based solutions.

for any s, discretized cell i, and discretized time step m. Since the hydraulic conductivity function is positive and symmetric along $x = 1$ m and $y = 1$ m, and $\nabla\psi\big|_{\omega^+ \triangleq [0,1]\times[0,1]\times z} = -\nabla\psi\big|_{\omega^- \triangleq [1,2]\times[1,2]\times z}$, we have $\sum_{j\in\mathcal{N}_i} \left[K(\Theta(\psi))\nabla(\psi+z)\right]^{m+1,s}_{\omega_{i,j}} \cdot \mathbf{n}_{\omega_{i,j}} A_{\omega_{i,j}} > 0$. Meanwhile, $\partial_t \Theta_i^{m+1}(\psi)\mathrm{vol}(V_i)$ is typically small due to the slow dynamics of water infiltration in soil and the fact that $\mathrm{vol}(V_i)$ is small. Thus, we have $\psi_{\text{analytical}} - \psi_{\text{numerical}} > 0$ for GRW-based solution, which explains the negativity of the relative error. On the other hand, for our D-GRW method, the use of MLPs to approximate the nonlinear mappings f and f^{-1} complicates the behavior (including the sign) of the relative error.

Regarding the distribution of the magnitude of relative error in the GRW solver, since hydraulic conductivity function is an increasing function of ψ, and ψ is at its maximum at the center of the x–y plain, it is expected that $\sum_{j\in\mathcal{N}_i} \left[K(\psi)\nabla(\psi+z)\right]^{m+1,s}_{\omega_{i,j}} \cdot \mathbf{n}_{\omega_{i,j}} A_{\omega_{i,j}}$, and hence the relative error, is maximized at and around the center of the x–y plane. However, for the D-GRW based pressure head solution, we suspect that the higher relative error at the four corners may be attributed to the slight decrease in accuracy of MLPs in approximating f and f^{-1} near the boundaries of domain.

Finally, we evaluate the MSE by summing the individual MSE values over all cells at $z = 0.5$ m and $z = 1$ m. For $z = 0.5$ m, $\mathrm{MSE}_{\text{D-GRW}}$ and $\mathrm{MSE}_{\text{GRW}}$ are calculated to be 8.044×10^{-4} and 0.1972, respectively. For $z = 1$ m, $\mathrm{MSE}_{\text{D-GRW}}$ and $\mathrm{MSE}_{\text{GRW}}$ are 6.736×10^{-3} and 0.5052, respectively. This indicates that the MSE of the GRW based pressure head solution is typically 1 to 2 orders of magnitude higher than our D-GRW based solution.

6 Conclusion

In this work, we present a novel data-facilitated numerical method for accurately and efficiently solving a general, d-dimensional Richards equation ($d = 1, 2, 3$). The numerical method adopts an adaptive linearization scheme based on the FVM discretization of the Richards equation to improve convergence and stability of the solution process. We can then solve the adaptive L-scheme formulation iteratively in a data-facilitated GRW model, in which we train two MLPs to approximate the nonlinear mapping and inverse mapping between pressure head and the number of particles in each cell. Additionally, we discuss effective ways to perform data augmentation, which can significantly improve the solution accuracy of our D-GRW method. Using three case studies, we illustrate that our numerical method not only achieves significantly improved solution accuracy, but also better preserves the overall mass balance and conservation laws compared to existing numerical methods. Furthermore, we believe that the proposed numerical framework integrating FVM, adaptive L-scheme, global random walk, and neural network approximation can be extended to effectively solving various transport equations with strong physical background and significance, such as the fractional diffusion model [47], water-soluble contaminant transport in soil, and so on.

7 Acknowledgments

We acknowledge financial support from the startup fund of College of Engineering, Architecture, and Technology at Oklahoma State University.

8 Data Availability Statement

We use the open-source platform SimPEG [48] and the code available at `https://simpegdocs.appspot.com/content/examples/20-published/plot_richards_celia1990.html` to implement the finite difference method (FDM) by Celia et al. [11] for the 1-D case study. The ground truth solutions shown in Figures 1, 3, and 4 are extracted from Celia et al. [11]. The GRW solver is adopted from [28], and the code is available at `https://github.com/PMFlow/FlowBenchmark`. The reference solution data generated for MLP training for all case studies are available at [49] in Zenodo. For the 2-D case study, we use HYDRUS 2D/3D Pro software (version 2.04.0580) in Figures 5 through 9. All figures are produced using OriginPro 2023b.

References

[1] E. Babaeian, M. Sadeghi, S. B. Jones, C. Montzka, H. Vereecken, M. Tuller, Ground, proximal, and satellite remote sensing of soil moisture, Reviews of Geophysics 57 (2) (2019) 530–616.

[2] D. Spelman, K.-D. Kinzli, T. Kunberger, Calibration of the 10HS soil moisture sensor for Southwest Florida agricultural soils, Journal of Irrigation and Drainage Engineering 139 (12) (2013) 965–971. `doi:10.1061/(ASCE)IR.1943-4774.0000647`.

[3] M. Saavoss, J. Majsztrik, B. Belayneh, J. Lea-Cox, E. Lichtenberg, Yield, quality and profitability of sensor-controlled irrigation: a case study of snapdragon (Antirrhinum majus L.) production, Irrigation Science 34 (5) (2016) 409–420.

[4] L. A. Richards, Capillary conduction of liquids through porous mediums, Physics 1 (5) (1931) 318–333.

[5] R. Haverkamp, M. Vauclin, J. Touma, P. Wierenga, G. Vachaud, A comparison of numerical simulation models for one-dimensional infiltration, Soil Science Society of America Journal 41 (2) (1977) 285–294.

[6] Y. Mualem, A new model for predicting the hydraulic conductivity of unsaturated porous media, Water Resources Research 12 (3) (1976) 513–522.

[7] M. T. Van Genuchten, A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Science Society of America Journal 44 (5) (1980) 892–898.

[8] W. Gardner, Some steady-state solutions of the unsaturated moisture flow equation with application to evaporation from a water table, Soil Science 85 (4) (1958) 228–232.

[9] M. W. Farthing, F. L. Ogden, Numerical solution of Richards' equation: A review of advances and challenges, Soil Science Society of America Journal 81 (6) (2017) 1257–1269.

[10] P. R. Day, J. N. Luthin, A numerical solution of the differential equation of flow for a vertical drainage problem, Soil Science Society of America Journal 20 (4) (1956) 443–447.

[11] 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.

[12] A. M. Ireson, R. J. Spiteri, M. P. Clark, S. A. Mathias, A simple, efficient, mass-conservative approach to solving Richards' equation (openre, v1. 0), Geoscientific Model Development 16 (2) (2023) 659–677.

[13] R. Eymard, M. Gutnic, D. Hilhorst, The finite volume method for Richards equation, Computational Geosciences 3 (3) (1999) 259–294.

[14] W. Lai, F. L. Ogden, A mass-conservative finite volume predictor–corrector solution of the 1D Richards' equation, Journal of Hydrology 523 (2015) 119–127.

[15] O. Misiats, K. Lipnikov, Second-order accurate monotone finite volume scheme for Richards' equation, Journal of Computational Physics 239 (2013) 123–137.

[16] S. Bassetto, C. Cancès, G. Enchéry, Q.-H. Tran, On several numerical strategies to solve Richards' equation in heterogeneous media with finite volumes, Computational Geosciences 26 (5) (2022) 1297–1322.

[17] D. Caviedes-Voullième, P. Garcı, J. Murillo, et al., Verification, conservation, stability and efficiency of a finite volume method for the 1D Richards equation, Journal of Hydrology 480 (2013) 69–84.

[18] G. Manzini, S. Ferraris, Mass-conservative finite volume methods on 2-D unstructured grids for the Richards' equation, Advances in Water Resources 27 (12) (2004) 1199–1215.

[19] I. S. Pop, F. Radu, P. Knabner, Mixed finite elements for the Richards' equation: linearization procedure, Journal of Computational and Applied Mathematics 168 (1-2) (2004) 365–373.

[20] L. Bergamaschi, M. Putti, Mixed finite elements and Newton-type linearizations for the solution of Richards' equation, International Journal for Numerical Methods in Engineering 45 (8) (1999) 1025–1046.

[21] K. Mitra, I. S. Pop, A modified L-scheme to solve nonlinear diffusion problems, Computers & Mathematics with Applications 77 (6) (2019) 1722–1738.

[22] G. Albuja, A. I. Ávila, A family of new globally convergent linearization schemes for solving Richards' equation, Applied Numerical Mathematics 159 (2021) 281–296.

[23] V. Casulli, P. Zanolli, A nested Newton-type algorithm for finite volume methods solving Richards' equation in mixed form, SIAM Journal on Scientific Computing 32 (4) (2010) 2255–2273.

[24] V. Zaburdaev, S. Denisov, J. Klafter, Lévy walks, Reviews of Modern Physics 87 (2) (2015) 483.

[25] A. Zoia, M.-C. Néel, A. Cortis, Continuous-time random-walk model of transport in variably saturated heterogeneous porous media, Physical Review E 81 (3) (2010) 031104.

[26] P. Dostert, Y. Efendiev, B. Mohanty, Efficient uncertainty quantification techniques in inverse problems for Richards' equation using coarse-scale simulation models, Advances in Water Resources 32 (3) (2009) 329–339.

[27] N. Suciu, Diffusion in Random Fields: Applications to Transport in Groundwater, Springer, 2019.

[28] N. Suciu, D. Illiano, A. Prechtel, F. A. Radu, Global random walk solvers for fully coupled flow and transport in saturated/unsaturated porous media, Advances in Water Resources 152 (2021) 103935.

[29] N. Suciu, C. Vamoş, I. Turcu, C. Pop, L. Ciortea, Global random walk modelling of transport in complex systems, Computing and Visualization in Science 12 (2) (2009) 77–85.

[30] C. Vamos, N. Suciu, H. Vereecken, O. Nitzsche, H. Hardelauf, Global random walk simulations of diffusion, in: Scientific Computing, Validated Numerics, Interval Methods, Springer, 2001, pp. 343–354.

[31] Z. Song, Z. Jiang, A data-driven random walk approach for solving water flow dynamics in soil systems, in: Proceedings of Foundations of Computer-Aided Process Operations and Chemical Process Control (FOCAPO/CPC), 2023, pp. 1–6.

[32] Z. Song, Z. Jiang, A data-driven modeling approach for water flow dynamics in soil, in: A. C. Kokossis, M. C. Georgiadis, E. Pistikopoulos (Eds.), 33rd European Symposium on Computer Aided Process Engineering, Vol. 52 of Computer Aided Chemical Engineering, Elsevier, 2023, pp. 819–824. doi:https://doi.org/10.1016/B978-0-443-15274-0.50131-1.
URL https://www.sciencedirect.com/science/article/pii/B9780443152740501311

[33] V. S. Sizikov, et al., Well-posed, ill-posed, and intermediate problems with applications, De Gruyter, 2011.

[34] L. C. Evans, Partial differential equations, Vol. 19, American Mathematical Soc., 2010.

[35] N. Abdellatif, C. Bernardi, M. Touihri, D. Yakoubi, A priori error analysis of the implicit Euler, spectral discretization of a nonlinear equation for a flow in a partially saturated porous media, Advances in Pure and Applied Mathematics 9 (1) (2018) 1–27.

[36] I. Pop, W. Yong, On the existence and uniqueness of a solution for an elliptic problem, Studia Univ. Babes-Bolyai Math 45 (2000) 97–107.

[37] H. Berninger, R. Kornhuber, O. Sander, Fast and robust numerical solution of the Richards equation in homogeneous soil, SIAM Journal on Numerical Analysis 49 (6) (2011) 2576–2597.

[38] F. Otto, L1-contraction and uniqueness for quasilinear elliptic–parabolic equations, Journal of Differential Equations 131 (1) (1996) 20–38.

[39] K. Hornik, Approximation capabilities of multilayer feedforward networks, Neural Networks 4 (2) (1991) 251–257.

[40] A. Pinkus, Approximation theory of the MLP model in neural networks, Acta Numerica 8 (1999) 143–195. doi:10.1017/S0962492900002919.

[41] I. B. V. Da Silva, P. J. Adeodato, PCA and Gaussian noise in MLP neural network training improve generalization in problems with small and unbalanced data sets, in: The 2011 International Joint Conference on Neural Networks, IEEE, 2011, pp. 2664–2669.

[42] D. Gąsiorowski, T. Kolerski, Numerical solution of the two-dimensional Richards equation using alternate splitting methods for dimensional decomposition, Water 12 (6) (2020) 1780.

[43] F. T. Tracy, Clean two-and three-dimensional analytical solutions of Richards' equation for testing numerical solvers, Water Resources Research 42 (8) (2006).

[44] A. L. Maas, A. Y. Hannun, A. Y. Ng, Rectifier nonlinearities improve neural network acoustic models, in: Proc. ICML, Vol. 30, Atlanta, Georgia, USA, 2013, p. 3.

[45] R. F. Carsel, R. S. Parrish, Developing joint probability distributions of soil water retention characteristics, Water Resources Research 24 (5) (1988) 755–769.

[46] J. Šimůnek, M. T. Van Genuchten, M. Šejna, Recent developments and applications of the HYDRUS computer software packages, Vadose Zone Journal 15 (7) (2016).

[47] E. Gerolymatou, I. Vardoulakis, R. Hilfer, Modelling infiltration by means of a nonlinear fractional diffusion model, Journal of Physics D: Applied Physics 39 (18) (2006) 4104.

[48] R. Cockett, S. Kang, L. J. Heagy, A. Pidlisecky, D. W. Oldenburg, Simpeg: An open source framework for simulation and gradient based parameter estimation in geophysical applications, Computers & Geosciences (2015).

[49] Z. Song, Z. Jiang, A data-facilitated numerical method for richards equation to model water flow dynamics in soil [Dataset], Zenodo (2023). doi:10.5281/zenodo.8400791.