Dissertation

submitted to the

Combined Faculty of Natural Sciences and Mathematics

of Heidelberg University, Germany

for the degree of

Doctor of Natural Sciences

Put forward by

Lukas Riedel

Born in: Speyer, Germany

Oral examination: 08 February 2021

Numerical Modeling and Data Assimilation of Soil Water Flow

Referees: Prof. Dr. Kurt Roth
Prof. Dr. Peter Bastian

Numerical Modeling and Data Assimilation of Soil Water Flow
Soil water flow is a key process in Earth's hydrological cycle and an essential part of many ecosystem services. Soils are porous media and exhibit a heterogeneous, multi-scale architecture. Their non-linear material properties have a significant influence on the soil water dynamics, which poses difficulties for numerical models. These material properties cannot be measured directly, but data assimilation methods can estimate them by combining information from measurements of soil hydraulic states and from numerical models. The validity of the estimation results can be strongly affected by model errors. This dissertation (i) presents a versatile software package for modeling soil water flow and analyzes the accuracy and efficiency of its numerical discretization schemes, and (ii) employs this software in synthetic data assimilation tasks to investigate the effects of unrepresented dynamics, topography, and small-scale heterogeneity on estimated material properties and forecasts conducted with them. The results reveal that favoring low-order numerical methods over more accurate ones can be justified for use cases in soil hydrology. Moreover, the findings indicate that one-dimensional models with estimated effective material properties can reasonably replicate the dynamics of heterogeneous, two-dimensional domains with complicated topography, if boundary conditions are represented correctly.

Numerische Modellierung und Datenassimilation von Bodenwasserfluss
Bodenwasserfluss ist ein Schlüsselprozess im Wasserkreislauf der Erde und ein wesentlicher Bestandteil vieler Ökosystemdienstleistungen. Böden sind poröse Medien und weisen eine heterogene Multiskalen-Architektur auf. Ihre nichtlinearen Materialeigenschaften haben einen signifikanten Einfluss auf die Bodenwasserdynamik, was Schwierigkeiten für numerische Modelle darstellt. Diese Materialeigenschaften können nicht direkt gemessen, aber mit Datenassimilationsmethoden geschätzt werden, indem Informationen aus Messungen bodenhydraulischer Zustände und aus numerischen Modellen kombiniert werden. Die Gültigkeit der Schätzungsergebnisse kann durch Modellfehler stark beeinträchtigt werden. Diese Dissertation (i) stellt ein vielseitiges Softwarepaket zur Modellierung des Bodenwasserflusses vor und analysiert die Genauigkeit und Effizienz seiner numerischen Diskretisierungsschemata, und (ii) setzt diese Software bei synthetischen Datenassimilationsaufgaben ein, um die Auswirkungen von nicht repräsentierter Dynamik, Topographie und kleinskaliger Heterogenität auf geschätzte Materialeigenschaften und damit durchgeführte Vorhersagen zu untersuchen. Die Ergebnisse zeigen, dass es für Anwendungsfälle in der Bodenhydrologie gerechtfertigt sein kann, numerische Methoden niedriger Ordnung genaueren Methoden vorzuziehen. Darüber hinaus weisen die Ergebnisse darauf hin, dass eindimensionale Modelle mit geschätzten effektiven Materialeigenschaften die Dynamik heterogener, zweidimensionaler Gebiete mit komplizierter Topographie angemessen nachbilden können, wenn die Randbedingungen korrekt repräsentiert werden.

Für meinen Vater,
in der Hoffnung, die Möglichkeiten
– oder zumindest einen Teil davon –
entdeckt zu haben.

Contents

List of Acronyms

AHFR average horizontal flux ratio

API application programming interface

DG discontinuous Galerkin

DOF degree of freedom

DORiE DUNE Operated Richards Equation Solving Environment

DUNE Distributed and Unified Numerics Environment

EnKF ensemble Kalman filter

FV finite volume

KnoFu Knowledge Fusion

MAV maximum averaging volume

NRMSE normalized root mean squared error

NSE Nash-Sutcliffe efficiency

ODE ordinary differential equation

PDE partial differential equation

pdf probability density function

REV representative elementary volume

RMSE root mean squared error

SWIP symmetric weighted interior penalty

TDR time domain reflectometry

VTK Visualization Toolkit

Introduction

Water is a key resource for humankind and Earth's hydrosphere is characterized by a continuous circulation of water (Oki and Kanae 2006). Soils are an essential compartment of this hydrological cycle, and their role is crucial for nearly all terrestrial systems. Soil water movement in the topmost, usually unsaturated layers of the soil is the interfacing process between atmospherical and lithospherical water flow. It supports several important ecosystem services, like retention, buffering, and filtering of fresh water. With the increasing anthropogenic influence on all of Earth's systems, soil science evolved along the technical possibilities to tackle research questions and the societal needs for the understanding of processes into a highly interdisciplinary field (Sivapalan and Blöschl 2017). Soils are the site of numerous physical, biological, and chemical processes involving fluxes not only of water, but also gases, solutes, and energy. Soil water content or moisture, however, is a key variable in most of these processes (Seneviratne et al. 2010).

Water flow through the porous medium soil is described by a physical model at the spatial scale of centimeters to meters (Roth 2012). General solutions to the highly non-linear equations require sophisticated numerical schemes (Farthing and Ogden 2017). They are further aggravated by the inherent multi-scale architecture of soils, which involves complicated geometries and heterogeneity at every spatial scale (Pachepsky and Hill 2017). The accuracy of models therefore hinges on the description of not only initial and boundary conditions, but especially material properties, which are usually expressed in parameterizations. Up- and downscaling from a specific spatial and temporal scale is only possible to a limited extent, as every scale features its own effective properties which are the result of dynamics operating on a lower scale (Vogel and Roth 2003).

Discontinuous Galerkin (DG) methods are both flexible and highly accurate numerical discretization schemes (Di Pietro and Ern 2012). Advances in computer architectures over the last decades have increased their popularity, as they feature a high potential for parallelization and the possibility to locally adjust their effective resolution. DG schemes can also be applied on complicated geometries. These features make them ideal candidates for modeling soil water flow, and formulations for heterogeneous convection-diffusion-reaction problems have been rigorously studied (Ern, Stephansen, and Zunino 2009). However, high accuracy and flexibility entail a complicated scheme formulation and a high computational cost. The selection of a numerical scheme therefore depends on the particular use case.

Hydrological processes like farmland irrigation, precipitation runoff, and groundwater uptake happen at the scale of fields, catchments, and landscapes. But the critical scale for the description of soil water movement is that of pedons (Vogel 2019). The intrinsic scaling problem forces scientists into a tradeoff between numerical accuracy and spatial resolution, which, coupled with the complicated soil architecture, has a profound influence on the results. In soil hydrology, simple and therefore computationally cheap discretization schemes at high resolutions are typically favored over more accurate methods (Miller et al. 2013). Large-scale models of catchments or landscapes often assume soil water flow to be mainly or exclusively one-dimensional, along the vertical axis (Zeng and Decker 2009).

The soil architecture and associated material properties are the largest source of uncertainty when modeling soil water flow. Soil hydraulic parameters are difficult to observe and cannot be inferred from laboratory experiments with soil samples. Material properties are therefore estimated by combining in situ measurements of system states with inverse models or data assimilation algorithms (Vrugt et al. 2008). A successful data assimilation algorithms in soil hydrology is the ensemble Kalman filter (EnKF; Evensen 2003), which has been used to estimate soil hydraulic states, parameters, and even architecture (Liu et al. 2012).

Measurements are usually scarce and noisy, and the model used for inversion or data assimilation has a critical role, as it must accurately recapture the physical processes observed. With the multi-scale nature of soils and the large number of processes involved, model representation errors are hardly avoidable. The focus in data assimilation therefore shifted from an estimation of the most likely states and parameters to a quantification of uncertainty in the estimation itself (Reichle 2008; Liu and Gupta 2007). The scientific community has identified the quantification of model errors as major challenge in soil hydrology, and recently developed several techniques to tackle it. These involve, for instance, using expert knowledge to identify unrepresented dynamics and to change the treatment of associated observations (Bauser et al. 2016), and estimating model errors from measurements alone with data driven methods (Zhang et al. 2019). Nonetheless, it remains an open question how errors in the representation of soil heterogeneity and architecture affect data assimilation results.

In my work and this thesis, I chose a twofold approach to the recent challenges in soil physics. With the DUNE Operated Richards Equation Solving Environment (DORiE; Riedel et al. 2020a), I co-authored a software package for modeling soil water flow and passive solute transport based on the Distributed and Unified Numerics Environment (DUNE; Bastian et al. 2020). It offers low accuracy finite volume (FV) and high accuracy DG discretization schemes for both processes and makes state-of-the-art numerical tools available to soil scientists without strong numerical background. Within this thesis, I showcase DORiE's features and investigate the performance of its discretization schemes with respect to applications in soil hydrology. I further developed the coupling between DORiE and a data assimilation framework and thus employed DORiE as forward model in an EnKF to research

the effect of unrepresented model errors. For a study I co-authored, we conducted synthetic experiments in which we used a one-dimensional model of a soil profile to estimate effective material properties with an EnKF based on measurements from a two-dimensional model with known small-scale heterogeneity (Bauser et al. 2020). In this thesis, I recapture the findings of this study and extend its analysis. I furthermore conduct additional synthetic experiments involving both unrepresented heterogeneity and topography in the estimation model, and investigate their effects onto the results.

Thesis Outline Chapter 1 introduces the physics of soil water flow. Chapter 2 derives the FV and DG discretization schemes for the physical equations and outlines the techniques required for computing solutions from them. The discussed numerical methods are implemented in DORiE, which is presented in chapter 3. This chapter also includes the analysis of its performance and accuracy. Chapter 4 introduces the EnKF, along with its extensions for applications in soil hydrology, and outlines the coupling between DORiE and an implementation of the filter. The study by Bauser et al. (2020) is revisited in chapter 5, with an expanded analysis of the results. In chapter 6, I extend these investigations by introducing unrepresented heterogeneity. Finally, this chapter also includes a combined conclusion from all conducted synthetic experiments involving the effects of unrepresented model errors.

Part I

Modeling Soil Hydrology

1 Soil Water Flow

Soil water flow is a central soil process and described with a continuum formulation of fluid flow in unsaturated, porous media (Vereecken et al. 2016). This formulation combines the laminar flow at the pore scale of micrometers to millimeters with a statistical description of the porous medium configuration. Crucial components of this theory are the observation that material properties of soils can be statistically averaged at certain length scales (Vogel 2019), and the assumption that the microscopic flow within the pores is in equilibrium with respect to the local pressure gradient. The porous medium is also assumed to be static.

All of these assumptions are observed to be frequently violated in real soils (Jury et al. 2011). Porous media typically exhibit hysteresis in the soil hydraulic state and intense forcings can break the locally equilibrated flow. This allows water to propagate along stable fingers and much faster than expected from the equilibrium model. Additionally, soils feature extended macropores, which are the result of geomorphological or biological activity. These facilitate possibly non-laminar flow processes with which water is able to bypass parts of, or the entire, porous matrix. Such processes are subsumed as preferential flow (Hendrickx and Flury 2001). In extreme cases, they can make up the main contribution to the overall soil water flow. Although numerous models for preferential flow exist, there is little consensus in their formulation and especially their coupling to the flow in the porous matrix (Šimůnek et al. 2003). In the following, I reiterate the derivation of the Richards equation, which neglects any of these effects but has proven robust enough to describe the dynamics of soil water flow in many real-world scenarios.

1.1 Capillary Water Flow

Assuming water to be an isothermal, incompressible, Newtonian fluid, its flow is described by the Navier-Stokes equation,

$$\rho \partial_t \boldsymbol{v} + \rho \left[\boldsymbol{v} \cdot \nabla\right] \boldsymbol{v} = \rho \boldsymbol{g} - \nabla p + \mu \nabla^2 \boldsymbol{v}, \tag{1.1}$$

with the water velocity $\boldsymbol{v}$ and pressure p, where ρ and μ are the water density and dynamic viscosity at constant temperature, respectively, and $\boldsymbol{g}$ indicates the gravitational acceleration. Analyzing the orders of magnitude of the constituents of the formula, we can approximate the flow in soil capillaries (Roth 2008). With a small Strouhal number, the forcing becomes

much slower than the resulting internal dynamics, and the flow is essentially stationary. We call this assumption *local equilibrium*. The Reynolds number is typically much smaller than the critical Reynolds number, indicating that dissipation is efficient and turbulence can be neglected. Equation (1.1) then simplifies to the stationary Stokes equation,

$$\mu \nabla^2 \boldsymbol{v} = \nabla p - \rho \boldsymbol{g}. \tag{1.2}$$

Pressure and gravity are the driving forces in Stokes flow, and in the stationary case, they are balanced by friction. As the only opposing force, the friction acts in the opposite direction of the flow. We use that eq. (1.2) is linear in $\boldsymbol{v}$ and insert a proportional relationship between friction and velocity, $\boldsymbol{v} = -\kappa \nabla^2 \boldsymbol{v}$, arriving at the simplified Stokes equation,

$$\boldsymbol{v} = -\frac{\kappa}{\mu} \left[\nabla p - \rho \boldsymbol{g} \right], \tag{1.3}$$

where $\kappa(\boldsymbol{x})$ represents the local geometry.

1.2 State Variables

Defining macroscopic state variables from microscopic quantities necessarily requires some form of spatial averaging, which is only feasible in a certain spatial range. The representative elementary volume (REV) defines the lower end of this range. It specifies the minimal volume for which the value of an averaged quantity becomes insensitive to small changes of said volume (Roth 2012). The upper end of the range is the maximum averaging volume (MAV), indicating the maximum extension for which the local equilibrium assumption is valid. Both definitions strongly depend on the microscopic system architecture, the current dynamics, and the external forcing. In general, however, there is no guarantee that the MAV is actually larger than the REV.

In spatial scales above the REV, we can then define the local *water content* θ as fraction of the water volume V_w in the total volume V,

$$\theta \coloneqq \frac{V_w}{V}. \tag{1.4}$$

The natural upper limit of θ is the saturated water content θ_s and equals the soil porosity ϕ, $\theta_s \coloneqq \phi$. The lower limit, the residual water content θ_r, is the fraction of immobile water remaining as adsorbed films in an otherwise dry medium on the timescale of interest. It is convenient to express the soil water content in terms of the soil water *saturation* $\Theta \in [0, 1]$,

$$\Theta \coloneqq \frac{\theta - \theta_r}{\theta_s - \theta_r}. \tag{1.5}$$

Further, we replace the description of force acting on a unit surface of the fluid by the potential energy density of a fluid unit volume. This matric potential ψ_m describes the pressure jump between water and air at their interface, $\psi_m = p_w - p_a$, and can be defined with the Young-Laplace equation,

$$\psi_m \coloneqq 2\sigma_{wa}H, \tag{1.6}$$

where σ_{wa} is the surface tension of the interface and H its mean curvature. The matric potential is often expressed as energy density per unit weight, yielding the *matric head*,

$$h_m \coloneqq \frac{\psi_m}{\rho g}, \tag{1.7}$$

where g is the gravitational acceleration. The value of the matric head is interpreted as height of an equivalent fluid column, $[h_m] = \mathrm{m}$.

1.3 Richards Equation

To derive the continuum formulation of the Richards equation from the obtained state variables and the pore-scale flow given by the Stokes equation, we further follow Roth (2008). The water flux at the macroscopic continuum scale is given by the water content θ multiplied with the spatial average of the flow velocity in the soil capillaries,

$$\boldsymbol{j} = \theta\langle \boldsymbol{v}^\mu \rangle_w = -\theta\frac{1}{\mu}\langle \kappa^\mu \left[\nabla p^\mu - \rho \boldsymbol{g}\right]\rangle_w, \tag{1.8}$$

where we inserted eq. (1.3) as the pore-scale velocity $\boldsymbol{v}^\mu$, indicating microscopic quantities with superscript μ, and where $\langle\cdot\rangle_w$ denotes the spatial average over the water phase. The average poses some difficulties due to the correlations between the pressure gradient and the pore geometry. However, owing to the linearity of the Stokes equation, the magnitude of the macroscopic pressure gradient must be proportional to the magnitude of the microscopic one. We can therefore relate them with a symmetric, second rank tensor $\boldsymbol{A}(\boldsymbol{x})$ through $[\nabla p^\mu - \rho \boldsymbol{g}] = \boldsymbol{A}[\nabla p - \rho \boldsymbol{g}]$, yielding Darcy's law,

$$\boldsymbol{j} = -\frac{1}{\mu}\theta\langle \kappa^\mu \boldsymbol{A}\rangle_w \left[\nabla p^\mu - \rho \boldsymbol{g}\right] = -\frac{\boldsymbol{k}}{\mu}\left[\nabla p - \rho \boldsymbol{g}\right], \tag{1.9}$$

where $\boldsymbol{k} \coloneqq \theta\langle \kappa^\mu \boldsymbol{A}\rangle_w$ is the permeability. Neglecting hysteresis, we can express $\boldsymbol{k}$ as a function of the water content θ. This function is non-linear because variations in θ change the averaging volume in $\langle\cdot\rangle_w$. Further assuming that the phases of water and air decouple, and that the water flux therefore is unaffected by pressure gradients in the air phase, we

Table 1.1: Mualem–van Genuchten parameter sets for four soil types used in simulations throughout this thesis, as reported by Carsel and Parrish (1988).

Medium	$K_0/\mathrm{m\,s^{-1}}$	$\alpha/\mathrm{m^{-1}}$	n	τ	θ_r	θ_s
Loam	2.9×10^{-6}	-3.6	1.56	0.5	0.078	0.43
Loamy Sand	4.1×10^{-5}	-12.4	2.28	0.5	0.057	0.41
Sandy Loam	1.2×10^{-5}	-7.5	1.89	0.5	0.065	0.41
Silt Loam	1.3×10^{-6}	-2.0	1.41	0.5	0.067	0.45

replace the water pressure with the matric potential ψ_m and define the hydraulic conductivity $\boldsymbol{K}'(\theta; \boldsymbol{x}) \coloneqq \boldsymbol{k}/\mu$ to arrive at the Buckingham-Darcy law,

$$\boldsymbol{j} = -\boldsymbol{K}'(\theta)\,[\nabla\psi_m - \rho\boldsymbol{g}]\,. \tag{1.10}$$

Inserting this into the continuity equation and dividing by ρg yields the Richards equation,

$$\frac{\partial\theta}{\partial t} - \nabla\cdot[\boldsymbol{K}(\theta)\,[\nabla h_m - \hat{\boldsymbol{g}}]] = 0, \tag{1.11}$$

which includes a redefinition of the hydraulic conductivity $\boldsymbol{K} \coloneqq \boldsymbol{K}'/(\rho g)$, $[\boldsymbol{K}] = \mathrm{m\,s^{-1}}$, and where $\hat{\boldsymbol{g}}$ is the unit vector in the direction of gravitational acceleration.

The rigid assumptions leading to the above formulation of the Richards equation have a profound influence on its applicability (Roth 2012). First, it is only valid for an intermediate soil water saturation Θ because we assumed a negligible influence of the air phase and its pressure gradients. For low soil water contents, and a consequently low magnitude of the hydraulic conductivity $\boldsymbol{K}$, vapor transport has a significant influence on the overall water flux inside the soil, but is neglected in the equation. In contrast, air bubbles are typically trapped in the water phase for high soil water saturations, requiring a two-phase flow formulation. Secondly, and as mentioned above, the local equilibrium assumption is a key component of the continuum formulation, but hinges on many factors like the small-scale architecture of the porous medium, internal state, and external forcing. Thirdly, solving the Richards equation requires a closure of the initially independent state variables h_m and θ. However, this closure is a direct consequence of assuming local equilibrium and neglecting hysteresis (Roth 2008). And finally, owing to the simplification of the Stokes flow and the continuum representation, the Richards equation does not consider preferential flow phenomena, as outlined in the introduction to this chapter.

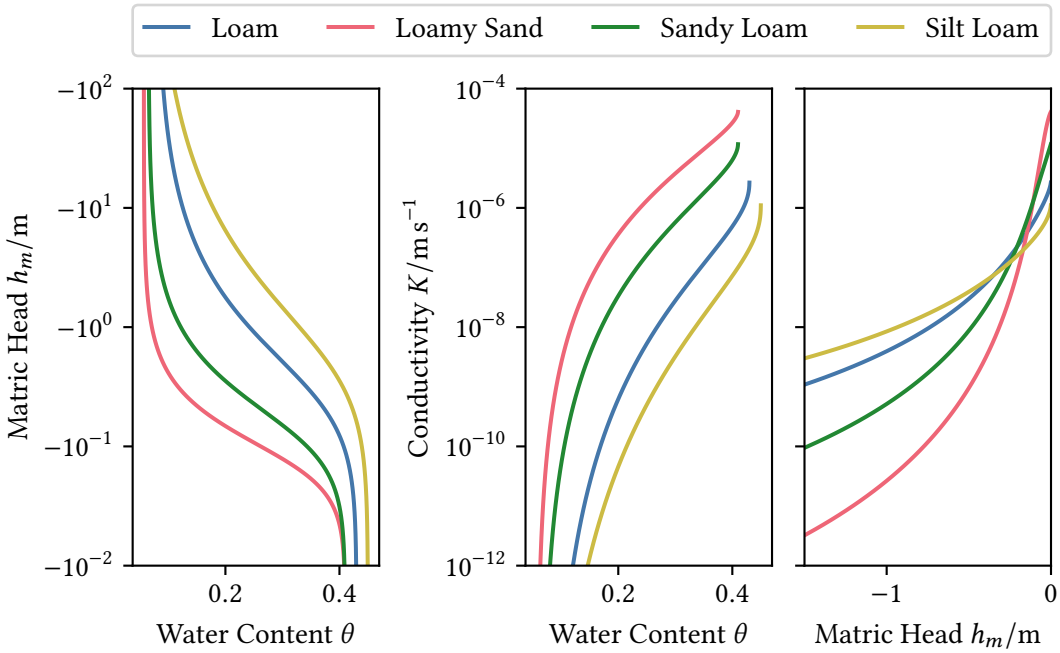

Figure 1.1: Retention curves (left) and conductivity functions (center, right) of parameter sets listed in table 1.1 according to the Mualem–van Genuchten parameterization. Note that for a positive matric head the conductivity equals the saturated conductivity, $K(h_m > 0\,\mathrm{m}) = K_0$.

1.4 Parameterization

To solve the Richards equation, eq. (1.11), we need to specify the hydraulic conductivity $\boldsymbol{K}(\theta)$ and the relation between matric head h_m and water content θ. These relations are parameterized into functions and describe the soil hydraulic properties of the particular medium. They encode the effect of these microscopic properties at the (macroscopic) continuum scale and therefore represent the subscale physics not explicitly considered in the Richards equation.

Assuming a completely wettable capillary with constant radius r in local hydraulic equilibrium, the capillary rise is given by the Young-Laplace equation, eq. (1.6), if we insert the radius as the inverse of the mean curvature, $H = r^{-1}$,

$$h_m = -\frac{1}{\rho g}\frac{\sigma_{wa}}{r}. \tag{1.12}$$

Further assuming that soil consists of several such capillaries with different radii, we can

therefore calculate the water content at any particular height using the above equation and a pore size distribution. However, this neglects any changes in the pore radii and overall architecture, and especially does not consider hysteresis. Presumably one of the most commonly used *soil water characteristics* or *retention curves* is the function defined by van Genuchten (1980),

$$\Theta(h_m) = \begin{cases} [1 + [\alpha h_m]^n]^{-1+1/n} & \text{if } h_m < 0, \\ 1 & \text{else,} \end{cases} \tag{1.13}$$

where n encodes the pore size distribution and α can be related to the inverse air entry value, $[\alpha] = \mathrm{m}^{-1}$. The air entry value h_a is the capillary rise of the largest pore with diameter $r_{\max}$, cf. eq. (1.12),

$$h_a = -\frac{1}{\rho g} \frac{\sigma_{wa}}{r_{\max}}. \tag{1.14}$$

In local equilibrium, the saturation drops below 1 at the height above the groundwater table which equals the air entry value.

The hydraulic conductivity $\boldsymbol{K}$ is typically considered to be isotropic, reducing the tensor to a scalar K. Mualem (1976) proposed a formulation assuming a proportional relation between capillary length and radius. Combining this with the retention curve by van Genuchten (1980) yields the isotropic hydraulic conductivity K,

$$K(\Theta) = K_0 \Theta^{\tau} \left[1 - \left[1 - \Theta^{n/(n-1)} \right]^{1-1/n} \right]^2, \tag{1.15}$$

with the saturated hydraulic conductivity K_0, $[K_0] = \mathrm{m\,s^{-1}}$ and the capillary tortuosity parameter τ. The soil hydraulic properties of a particular medium with respect to the Richards equation are then characterized by the six parameters used in eqs. (1.5), (1.13) and (1.15).

The lack of an explicit air entry value h_a in the van Genuchten saturation model, eq. (1.13), allows to compute the derivation of the parameterization functions for all potentials $h_m < 0\,\mathrm{m}$. But for fine-grained soils with $n < 2$, the conductivity function becomes increasingly non-linear towards full saturation, which is a major issue for numerical solvers. Vogel, van Genuchten, and Cislerova (2000) therefore propose to incorporate an explicit air entry value to reduce the non-linearity of the model. They demonstrate that this has profound influences on numerical results and stabilizes numerical simulations. The Mualem–van Genuchten model with explicit air entry value also produces simulation results that are more consistent with experimental data (Ippisch, Vogel, and Bastian 2006). However, with a lack of literature providing soil hydraulic properties in this modified model, I use the standard Mualem–van Genuchten model in numerical simulations throughout this thesis.

1.5 Heterogeneity and Miller Scaling

Soils are inherently heterogeneous media with observable layers, featuring sharp interfaces or smooth transitions between different soil types, depending on the spatial scale of interest (Vogel 2019). Heterogeneity is therefore generally modeled in two different ways; either by distinct parameter sets for the parameterization functions, or by scaling fields which manipulate these scaling functions.

If the transition from one medium to another is sufficiently sharp, the MAV can shrink below the spatial scale of interest. It is then suitable to characterize each medium with a different set of Mualem–van Genuchten parameters, creating a discontinuity of the parameterization functions at the medium interface. This breaks the notion of the REV and MAV across the interface itself, but not inside the respective media. Still, the local equilibrium assumption enforces a continuous potential, and the mass balance enforces a continuous flux, across the interface. Following eq. (1.10), the conductivity, water content, and the gradient of the matric head adjust to supply this flux. As the conductivity, water content, and matric head directly depend on each other through the material properties, the discontinuity is propagated to all related quantities except the matric head and water flux.

While the above approach constitutes large-scale heterogeneity by defining different soil layers with specific extensions, small-scale heterogeneity inside a single soil layer is modeled by locally modifying the parameterization functions. One possibility is Miller scaling, where a single spatial scaling factor ξ is used to scale the reference parameterization functions of the homogeneous medium (Miller and Miller 1956). The scaling factor is the ratio between the reference pore length scale $\ell^\dagger$ and the local pore length scale $\ell(\boldsymbol{x})$,

$$\xi(\boldsymbol{x}) \coloneqq \frac{\ell(\boldsymbol{x})}{\ell^\dagger}, \tag{1.16}$$

where $\boldsymbol{x}$ is the position. A scaling factor of $\xi > 1$ therefore creates a more coarse-grained, and a factor of $\xi < 1$ a more fine-grained material than the reference material. The reference parameterization functions denoted by superscript † are locally scaled according to

$$h_m(\cdot;\boldsymbol{x}) = \xi^{-1}(\boldsymbol{x}) \cdot h_m^\dagger(\cdot), \tag{1.17}$$

$$K(\cdot;\boldsymbol{x}) = \xi^2(\boldsymbol{x}) \cdot K^\dagger(\cdot). \tag{1.18}$$

This scaling assumes that all regions within a soil layer are geometrically similar at the pore scale. Consequently, the porosity is an invariant of the scaled medium. Within the framework of Miller scaling there are no restrictions on the nature of the field defining the local scaling factor ξ.

1.6 Boundary Conditions

There are two conventional ways of specifying the boundary conditions in the Richards regime. The canonical way is stating the value of the matric head h_m or the water content θ at the boundary. This is a Dirichlet problem, and through the soil water characteristic, eq. (1.13), all values at the boundary are consequently known. In a stationary setting, this problem yields a unique solution for the entire system.

The second way is stating the boundary flux, constituting a Neumann problem. As both the gradient of the matric head, and—through the conductivity function—the matric head itself contribute to the flux according to eq. (1.10), a Neumann problem alone yields no unique solution. Only when assuming a stationary setting in which there are no capillary forces acting on the water and thus $\nabla h_m = 0$, the flux is determined by the conductivity only,

$$\boldsymbol{j} = K(h_m)\hat{\boldsymbol{g}}. \tag{1.19}$$

The solution is then given by the matric head value for which the above equation holds. This *gravity flow* assumption is suitable at a sufficient distance from the water table.

Another typical boundary condition in soil hydraulic applications is seepage, which allows water to exit the domain without imposing a net force. This can be realized through a Dirichlet boundary condition with $h_m = 0\,\mathrm{m}$ (implying no net pressure at the water-air interface) that turns into a no-flow Neumann boundary condition as soon as this potential would lead to a flux into the domain (Scudeler et al. 2017). Additionally, a boundary condition of this type with a negative potential can also be used to model an evaporation boundary condition.

2 Numerical Solutions to the Richards Equation

In this chapter, I derive the FV and DG discretization schemes for the Richards equation. Cell-centered FV schemes are very simple discretization formulations. They only contain sums over grid faces, require no evaluation of basis functions, and avoid numerical integration. However, in their basic formulation, they are only applicable on structured, rectangular meshes. DG schemes, on the other hand, can be applied on any convex cell geometry and even irregular grids with hanging nodes. They also achieve better theoretical convergence rates, error estimates, and a higher accuracy than FV methods with the same number of degrees of freedom (DOFs) (Di Pietro and Ern 2012). However, they also require more computation time for simulations on the same mesh. In many cases, the grid resolution cannot be chosen arbitrarily but is limited through the structure of the simulated domain. If numerical accuracy is a minor concern, FV or low-order finite element schemes are therefore favored over DG schemes in soil hydrology (Miller et al. 2013).

The Richards equation, eq. (1.11), poses several problems for numerical discretization schemes and solvers. In the previously presented mixed form it appears as a diffusion equation, but both its diffusion and instationary term are non-linear because of the parameterization functions $K(\cdot)$ and $\theta(\cdot)$. It is therefore difficult to classify. The Richards equation is a parabolic partial differential equation (PDE) that becomes elliptic near saturation, and even effectively hyperbolic in certain scenarios with strong convection (Ippisch 2016). Additionally, the equation is also degenerate because of the highly non-linear parameterization and possibly non-existent derivatives. Finally, depending on the forcing and the material properties, the dynamics can exhibit large gradients in pressure, conductivity, and water content. These issues demand a careful numerical treatment and especially robust solvers (Farthing and Ogden 2017).

2.1 Discrete Setting and Grid Function Spaces

We compute the solutions on the physical domain $\Omega \subset \mathbb{R}^d$ with boundary $\partial\Omega$, where $d \in \{2, 3\}$ indicates the spatial dimensions. The domain is tessellated into a mesh $\mathcal{T}_h$ consisting of simplicial or rectangular grid elements T. All faces F inside the domain are collected in $\mathcal{F}_h$, with subsets for boundary faces, $\mathcal{F}_h^b \ni F \subset \partial\Omega$, and interfaces between adjacent grid elements, $\mathcal{F}_h^i \ni F = \partial T^+ \cap \partial T^-$. The domain boundary $\partial\Omega$ is divided into the segments Γ^N and Γ^D with respective boundary face subsets $\mathcal{F}_h^N, \mathcal{F}_h^D \subseteq \mathcal{F}_h^b$, on which

Neumann and Dirichlet boundary conditions are applied, respectively. Note that this division of the boundary may be variable in case of transient boundary conditions. For each face F a unit normal vector $\hat{\boldsymbol{n}}_F$ is defined pointing from T^+ towards T^- or coinciding with the outward unit normal vector on $\partial\Omega$ if $F \in \mathcal{F}_h^b$.

The discontinuous finite element function space of order k on $\mathcal{T}_h$ is given by

$$V_h = \left\{ v \in L^2(\Omega) : \forall T \in \mathcal{T}_h, v|_T \in \mathcal{P}_k \right\}, \tag{2.1}$$

where $L^2(\Omega)$ is the space of all square-integrable functions on Ω, and where $\mathcal{P}_k$ is the set of polynomials of total degree k, $\mathbb{P}_k$, on simplex grid elements and the set of polynomials of maximum degree k, $\mathbb{Q}_k$, on rectangular grid elements. Spaces $\mathcal{P}_k$ with degree $k > 0$ yield "broken" polynomial spaces used for DG discretization schemes. Polynomials of degree 0 on rectangular grid elements, $\mathbb{Q}_0$, yield a FV space with a single DOF, and consequently a constant value, per grid element.

Following the notation from above, we denote the restrictions $v|_{T^\pm}$ of any function v on grid elements $T^\pm$ by $v^\pm$ and define the jump at location $\boldsymbol{x} \in \partial T^+ \cap \partial T^-$,

$$[\![v]\!](\boldsymbol{x}) \coloneqq v^+(\boldsymbol{x}) - v^-(\boldsymbol{x}), \tag{2.2}$$

and the weighted average with positive weights $\omega^- + \omega^+ = 1$,

$$\{\!\{v\}\!\}_\omega(\boldsymbol{x}) \coloneqq \omega^+ v^+(\boldsymbol{x}) + \omega^- v^-(\boldsymbol{x}), \tag{2.3}$$

for all interface locations $\boldsymbol{x} \in F \in \mathcal{F}_h^i$. These definitions of jump and average are extended to locations on boundary faces $\boldsymbol{x} \in F \in \mathcal{F}_h^b$ by selecting the respective "interior" value,

$$[\![v]\!](\boldsymbol{x}) = \{\!\{v\}\!\}_\omega(\boldsymbol{x}) = v^+(\boldsymbol{x}). \tag{2.4}$$

We further denote the harmonic average of two variables a and b by $\langle a, b \rangle \coloneqq 2ab/(a+b)$. The inner product of two functions v and w over a domain $Q \subset \mathbb{R}^d$ is defined by

$$(v, w)_Q \coloneqq \int_Q vw \, \mathrm{d}x, \tag{2.5}$$

and used to compute the volume of said domain as $|Q| = (1, 1)_Q$.

2.2 Strong Problem Formulation

Given the above definitions, the strong problem of finding the solution $u(\boldsymbol{x}; t)$ of the matric head h_m for any location $\boldsymbol{x} \in \Omega$ and any time t in the closed time span $\Sigma = [t_0, t_0 + T]$ reads

$$\frac{\partial}{\partial t}\theta(h_m) - \nabla \cdot [K(h_m)\,[\nabla h_m - \hat{\boldsymbol{g}}]] = 0 \qquad \text{in } \Omega \times \Sigma, \tag{2.6a}$$

$$h_m = u_D(\boldsymbol{x}; t) \qquad \text{on } \partial\Omega_D, \tag{2.6b}$$

$$K(h_m)\,[\nabla h_m - \hat{\boldsymbol{g}}] \cdot \hat{\boldsymbol{n}}_\Omega = j_N(\boldsymbol{x}; t) \qquad \text{on } \partial\Omega_N, \tag{2.6c}$$

$$h_m(\boldsymbol{x}; t = t_0) = u_0(\boldsymbol{x}) \qquad \text{in } \Omega, \tag{2.6d}$$

where u_D and j_N are the Dirichlet and Neumann boundary conditions, respectively, u_0 is the initial condition, and $\hat{\boldsymbol{n}}_\Omega$ is the outward unit normal vector at the boundary of Ω. Note that with this definition of the Neumann flux, the direction of the flux is outward for positive values and inward for negative values of j_N.

We discretize the problem by applying the method of lines, which transforms it into a set of ordinary differential equations (ODEs). This will then be solved by a Runge-Kutta method. The weak problem formulation splits the forms for spatial semi-discretization, α_h and λ_h, from the temporal semi-discretization β_h,

$$\alpha_h(u_h(t), v_h; t) + \frac{\partial}{\partial t}\beta_h(u_h(t), v_h; t) = \lambda_h(v_h; t), \tag{2.7}$$

where $u_h(t) \in [V_h \times \Sigma]$ is the solution in the finite element function space, and $v_h \in V_h$ is the test function.

By convention, I will denote the numerical solution with u_h and identify it with the unknown variable, the matric head, $u_h \equiv h_m$. I consider both quantities synonymous in the context of numerical solvers, but will use u_h and h_m to emphasize numerical and physical considerations, respectively.

2.3 Finite Volume Discretization

FV methods use a $\mathbb{Q}_0$ function space with a single value of $u_h \in V_h$, and consequently a single hydraulic conductivity K and water content θ, defined per grid element T. At interfaces, the gradient is approximated by the jump of the solution $[[u_h]]$ divided by the distance of the adjacent grid element barycenters $\ell_{T^\pm}$. This assumes a structured rectangular mesh, where the normal flux through one side of a grid element completely describes the flow out of that grid element in this particular direction. The harmonic average of conductivities at the interface is a natural choice for approximating the interface conductivity, but can lead to instabilities when the dynamics become effectively convective (Ippisch 2016). We therefore

select an upwind matric head u_h^* for each interface F based on the direction of the flux,

$$u_h^* = \begin{cases} u_h^+ & \text{if } [[u_h]]/\ell_{T^\pm} - \hat{\boldsymbol{g}} \cdot \hat{\boldsymbol{n}}_F > 0, \\ u_h^- & \text{else,} \end{cases} \tag{2.8}$$

which is then used to evaluate the conductivity on either side,

$$K_*^\pm = K^\pm(u_h^*). \tag{2.9}$$

Dirichlet boundary conditions are applied weakly by redefining the jump of the solution $[[u_h]]$ as

$$[[u_h]] \coloneqq \begin{cases} u_h^+ - u_D & \text{if } F \in \mathcal{F}_h^D, \\ u_h^+ - u_h^- & \text{if } F \in \mathcal{F}_h^i, \end{cases} \tag{2.10}$$

where u_D is the boundary condition value at the respective boundary segment, as defined in eq. (2.6b), and by dividing this jump by the distance between grid cell barycenter and boundary $\ell_{T,F}$.

The resulting FV spatial discretization reads

$$\begin{aligned} \alpha_h^{\mathrm{FV}}(u_h) = &\sum_{F \in \mathcal{F}_h^i} \langle K_*^+, K_*^- \rangle \left[\frac{[[u_h]]}{\ell_{T^\pm}} - \hat{\boldsymbol{g}} \cdot \hat{\boldsymbol{n}}_F \right] |F| \\ &+ \sum_{F \in \mathcal{F}_h^D} K_*^+ \left[\frac{[[u_h]]}{\ell_{T,F}} - \hat{\boldsymbol{g}} \cdot \hat{\boldsymbol{n}}_F \right] |F|, \end{aligned} \tag{2.11}$$

where we dropped the dependence between hydraulic conductivity and matric head, $K = K(u_h)$, for brevity. This dependency also makes α_h non-linear in u_h. The Neumann boundary condition contributes to the constant form,

$$\lambda_h^{\mathrm{FV}} = -\sum_{F \in \mathcal{F}_h^N} j_N |F|, \tag{2.12}$$

and the temporal discretization is given by

$$\beta_h^{\mathrm{FV}}(u_h) = \sum_{T \in \mathcal{T}_h} \theta(u_h)|T|. \tag{2.13}$$

2.4 Discontinuous Galerkin Discretization

Diffusion dependent weights for interior penalty methods were analyzed by Dryja (2003) and later extended for heterogeneous advection diffusion reaction equations by Ern, Stephansen, and Zunino (2009). We use the symmetric weighted interior penalty (SWIP) DG formulation for heterogeneous diffusion by Di Pietro and Ern (2012, Chapter 4.5), which revisits both approaches. We apply their SWIP bilinear form to the stationary Richards equation, eq. (1.11) with $\partial_t \theta = 0$, by (i) inserting a function $u_h \in V_h$ of the broken polynomial space as solution, (ii) computing the scalar product with a test function $v_h \in V_h$, (iii) writing the integral as a sum over integrals on single grid elements T, (iv) integrating by parts, (v) formulating the resulting sum over mesh faces F instead of element boundaries ∂T, (vi) adding a penalty term for jumps in solution and test function to achieve coercivity, and (vii) adding a symmetry term, yielding

$$\alpha_h^{\mathrm{DG}}(u_h, v_h) = \sum_{T \in \mathcal{T}_h} \left(\nabla v_h, K\left[\nabla u_h - \hat{\boldsymbol{g}}\right]\right)_T \tag{2.14a}$$

$$- \sum_{F \in \mathcal{F}_h^i} \left([\![v_h]\!], \{\!\{K\left[\nabla u_h - \hat{\boldsymbol{g}}\right]\}\!\}_{\omega_K} \cdot \hat{\boldsymbol{n}}_F\right)_F \tag{2.14b}$$

$$- \sigma \sum_{F \in \mathcal{F}_h^i} \left([\![u_h]\!], \{\!\{K \nabla v_h\}\!\}_{\omega_K} \cdot \hat{\boldsymbol{n}}_F\right)_F \tag{2.14c}$$

$$+ \sum_{F \in \mathcal{F}_h^i} \frac{\eta}{h_F} \left([\![v_h]\!], \left\langle K^-, K^+ \right\rangle [\![u_h]\!]\right)_F \tag{2.14d}$$

$$+ \alpha_h^{\mathrm{DG},D}(u_h, v_h), \tag{2.14e}$$

where $\alpha_h^{\mathrm{DG},D}$ contains the Dirichlet boundary condition terms given in the next paragraph. The symmetry parameter $\sigma \in \{1, 0, -1\}$ can be used to set a symmetric, incomplete, or non-symmetric interior penalty formulation, respectively. In this work, I will only discuss the symmetric case, $\sigma = 1$, which simplifies solving the resulting linear system and leads to an optimal L^2-norm error estimate (Di Pietro and Ern 2012). We use the hydraulic conductivity K evaluated at either side of the interface as weights ω_K of the gradient averages,

$$\omega_K^{\pm} = \frac{K^{\mp}}{K^{\pm} + K^{\mp}}, \tag{2.15}$$

where we imply that the restriction of K to one grid element $T^{\pm}$ is extended to the solution with which it is evaluated, $K^{\pm} = K^{\pm}(h_m^{\pm})$. The local length scale h_F is given by the ratio between the smaller adjacent grid element volume and the interface volume,

$$h_F = \frac{\min(|T^+|, |T^-|)}{|F|}. \tag{2.16}$$

We follow Bastian (2014) in the definition of the penalty factor η, which includes the basis function polynomial order k, the spatial dimension d, and a user-defined "penalty factor" parameter m,

$$\eta = m \cdot p(p + d - 1). \tag{2.17}$$

The summand given by eq. (2.14d) is called *penalty term* and used to achieve discrete coercivity of the formulation, which is the case whenever η is sufficiently large. We choose $m = 10$. The temporal term of the Richards equation does not contain any derivatives, so it suffices to compute the scalar product,

$$\beta_h^{\mathrm{DG}}(u_h, v_h) = \sum_{T \in \mathcal{T}_h} (v_h, \theta(u_h))_T. \tag{2.18}$$

The Neumann boundary conditions constitute the only contribution to the linear form,

$$\lambda_h^{\mathrm{DG}}(v_h) = - \sum_{F \in \mathcal{F}_h^N} (v_h, j_N)_F. \tag{2.19}$$

Dirichlet boundary conditions are again incorporated weakly by modifying the definition of the solution jump at boundaries according to eq. (2.10). This can lead to macroscopic jumps $[\![u_h]\!]$, for which the same considerations apply as for jumps in the FV scheme. At Dirichlet boundaries, we select an upwind matric head u_h^* according to the direction of the numerical flux,

$$u_h^* = \begin{cases} u_h^+ & \text{if } \eta [\![u_h]\!] / h_F - [\nabla u_h - \hat{\boldsymbol{g}}] > 0, \\ u_h^- & \text{else.} \end{cases} \tag{2.20}$$

This estimation only neglects the harmonic average used in the penalty term, eq. (2.14d), and thus considers local gradient and jump equally. In cases where this assumption is erroneous, our particular choice favors the conductivity inside the domain against the conductivity induced by applying the boundary condition. The Dirichlet boundary condition terms then read

$$\begin{aligned} \alpha_h^{\mathrm{DG},D}(u_h, v_h) = &- \sum_{F \in \mathcal{F}_h^D} \left(v_h, K_*^+ [\nabla u_h - \hat{\boldsymbol{g}}] \cdot \hat{\boldsymbol{n}}_F\right)_F \\ &- \sigma \sum_{F \in \mathcal{F}_h^D} \left([\![u_h]\!], K_*^+ \nabla v_h \cdot \hat{\boldsymbol{n}}_F\right)_F \\ &+ \sum_{F \in \mathcal{F}_h^D} \frac{\eta}{h_F} \left(v_h, K_*^+ [\![u_h]\!]\right)_F. \end{aligned} \tag{2.21}$$

The treatment of the non-linearities introduced by the parameterization functions $\theta(\cdot)$ and $K(\cdot)$ is a prime concern when solving the equation. Through these non-linearities, the order of a polynomial representation of the function results is typically higher than the order of

the inserted solution u_h. In the spatial discretization, we aim at improving the sampling of these functions by increasing the order of numerical quadrature for the scalar products calculated in eqs. (2.14) and (2.18). With the polynomial order k of the solution $u_h \in V_h$, we employ a numerical quadrature of polynomial order $2 + 2k$ in the space semi-discretization α_h^{DG}, and a quadrature of order $2k$ in the time semi-discretization β_h^{DG}.

The SWIP DG formulation is well-posed and has a unique solution if K is a constant scalar per grid cell T (Di Pietro and Ern 2012), and also, in an extended form, if K is a tensor (Ern, Stephansen, and Zunino 2009). To the best of my knowledge, no rigorous mathematical analysis of the SWIP DG discretization with a conductivity or permeability depending on the solution, either in a linear or non-linear relationship, has been published yet. However, there are numerous successful applications of DG discretization schemes on the Richards equation published in literature (e.g. Li, Farthing, and Miller 2007; Solin and Kuraz 2011; Clément et al. 2020).

2.5 Solver Application

After selecting a tessellation $\mathcal{T}_h(\Omega)$, any function v_h of the broken polynomial function space $V_h(\Omega)$ can be expressed as linear combination of a set of suitable basis functions $\Phi = \{\varphi_i\}_{i=1,\dots,N_{\mathrm{DOF}}}$, where $N_{\mathrm{DOF}} = \dim V_h$, with independent coefficients $\boldsymbol{u} \in \mathbb{R}^{N_{\mathrm{DOF}}}$, which encode the DOFs. Additionally, we choose a finite set of points in time $\{t_k\}_{k=1,\dots,N_t}$ at which we want to compute the solution. We can then define two discrete operators for expressing the problem given by eq. (2.7),

$$[\mathcal{A}(\boldsymbol{u}_k; t_k)]_i \coloneqq \alpha_h^{\mathrm{m}}\Big(\sum_{j=1}^{N_{\mathrm{DOF}}} [\boldsymbol{u}_k]_j \, \varphi_j,\ \varphi_i;\ t_k \Big) - \lambda_h^{\mathrm{m}}\left(\varphi_i; t_k\right), \tag{2.22}$$

$$[\mathcal{B}(\boldsymbol{u}_k)]_i \coloneqq \beta_h^{\mathrm{m}}\Big(\sum_{j=1}^{N_{\mathrm{DOF}}} [\boldsymbol{u}_k]_j \, \varphi_j,\ \varphi_i \Big), \tag{2.23}$$

where $\mathrm{m} = \{\mathrm{FV}, \mathrm{DG}\}$ encodes the model type.

Runge-Kutta methods are one-step, multi-stage iterative methods for numerical integration. In the Shu-Osher representation, explicit and diagonally implicit Runge-Kutta methods with s stages are defined by two lower Hessenberg matrices $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^{s \times s+1}$ and a vector $\boldsymbol{\delta} \in \mathbb{R}^{s+1}$ (Ferracina and Spijker 2005). For every stage $i = 1, \dots, s$, the intermediate solution $\boldsymbol{u}_k^{(i)}$ is

found by solving

$$\sum_{j=0}^{i} \underbrace{\left[\frac{\alpha_{ij}}{t_{k+1} - t_k} \mathcal{B}\left(\boldsymbol{u}_k^{(j)}\right) - \beta_{ij} \mathcal{A}\left(\boldsymbol{u}_k^{(j)}, t_k^{(j)}\right) \right]}_{:= \mathcal{R}\left(\boldsymbol{u}_k^{(j)}\right)} = \emptyset, \tag{2.24}$$

where the intermediate time steps are given by

$$t_k^{(j)} = t_k + \delta_j (t_{k+1} - t_k). \tag{2.25}$$

The initial value of the intermediate solutions is the solution for time step k, and the solution of the final stage $i = s$ is the solution for the next time step $k + 1$,

$$\boldsymbol{u}_k = \boldsymbol{u}_k^{(0)}, \quad \boldsymbol{u}_{k+1} = \boldsymbol{u}_k^{(s)}. \tag{2.26}$$

Implicit Runge-Kutta methods are especially suitable for stiff PDEs because they are unconditionally stable with respect to the time step size. However, eq. (2.24) then results in a system of algebraic equations with $s \times N_{\mathrm{DOF}}$ components. We employ the two-stage, diagonally implicit Runge-Kutta method by Alexander (1977) with

$$\boldsymbol{\alpha} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{\beta} = \begin{pmatrix} 0 & \vartheta & 0 \\ 0 & 1 - \vartheta & \vartheta \end{pmatrix}, \quad \boldsymbol{\delta} = \begin{bmatrix} 0 \\ \vartheta \\ 1 \end{bmatrix}, \tag{2.27}$$

and $\vartheta = 1 - \sqrt{2}/2$, where the two stages are decoupled equation systems and can be solved successively.

We solve the non-linear, algebraic equation system given by eq. (2.24) using an iterative Newton method with line search algorithm. The solution is found by minimizing the residual $\mathcal{R}(\boldsymbol{u})$ with respect to the solution $\boldsymbol{u}$. The Newton method achieves quadratic convergence to a local minimum but requires the assembly of a Jacobian $\boldsymbol{J}(\boldsymbol{u})$,

$$J_{ij}(\boldsymbol{u}) = \frac{\partial}{\partial u_j} \left[\mathcal{R}(\boldsymbol{u})\right]_i, \tag{2.28}$$

which makes it computationally expensive. The application of different iteration schemes on the Richards equation has been studied extensively (Farthing and Ogden 2017). List and Radu (2016) find that mixed schemes with approximated Jacobians can achieve global convergence at a lower computational cost than a regular Newton scheme. However, they do not discuss Newton schemes with line search algorithms which increase the convergence radius. An analytical formulation of the Jacobian would require derivations of parameterization functions, which are not universal. We therefore use numerical differentiation to estimate

the Jacobian through

$$\boldsymbol{J}(\boldsymbol{u}) \approx \frac{\mathcal{R}(\boldsymbol{u}+\boldsymbol{\epsilon}) - \mathcal{R}(\boldsymbol{u})}{\|\boldsymbol{\epsilon}\|}, \tag{2.29}$$

with some small difference vector $\boldsymbol{\epsilon}$. This estimation increases the computational cost due to repeated applications of the residual operator $\mathcal{R}(\cdot)$. On the other hand, Bastian (2014) argues that a numerical estimation of the Jacobian "enables the treatment of nonsmooth nonlinearities," which may circumvent possible degeneracy of the parameterization functions and thus be advantageous to analytical computations.

Each step $n \geq 0$ of the Newton solver requires a solver for computing the update $\boldsymbol{z}_n$ from the resulting linear equation system,

$$\boldsymbol{J}_n \boldsymbol{z}_n - \boldsymbol{r}_n = \boldsymbol{0}, \tag{2.30}$$

with $\boldsymbol{J}_n = \boldsymbol{J}(\boldsymbol{u}_n)$ and $\boldsymbol{r}_n = \mathcal{R}(\boldsymbol{u}_n)$. The update is then applied to the solution through

$$\boldsymbol{u}_{n+1} = \boldsymbol{u}_n - \sigma \boldsymbol{z}_n, \quad \sigma \in (0, 1]. \tag{2.31}$$

For the DG discretization, we apply the algebraic multigrid solver presented by Bastian (2014). It splits the problem into a continuous Galerkin and a FV subspace and thus enables to apply different preconditioners for the "low frequency" and "high frequency" residual contributions. For the FV problem, we use a similar multigrid BiCGSTAB solver with SSOR preconditioner, but without specialization for DG discretizations. With an absolute limit $\epsilon_{\text{abs}} \ll 1$ and a reduction limit $\epsilon_{\text{red}} \ll 1$, the Newton solver is considered converged if $\|\boldsymbol{r}_n\| < \epsilon_{\text{abs}}$ or $\|\boldsymbol{r}_n\| / \|\boldsymbol{r}_0\| < \epsilon_{\text{red}}$.

2.6 Adaptive Time Step Scheme

Adapting time steps is a key ingredient for efficient and precise numerical time-stepping methods. Typical control algorithms adapt the time step size according to an error measure, which requires knowledge of the underlying problem formulation (Söderlind 2006). Implicit time step schemes are generally unconditionally stable with respect to the time step size. However, the locality of the Newton method and its limited capability of linearizing the equation system poses implicit restrictions on the time step size, which also depend on the current system state and forcing.

We use a heuristic approach based on the convergence behavior of the Newton method. Generally speaking, the fewer iterations the Newton solver requires to converge, the better is the linear approximation of the problem. A nearly linear problem in turn implies weak dynamics in the Richards regime or a very small time-step, where the current solution is already a good approximation for the next time step (Farthing and Ogden 2017). We

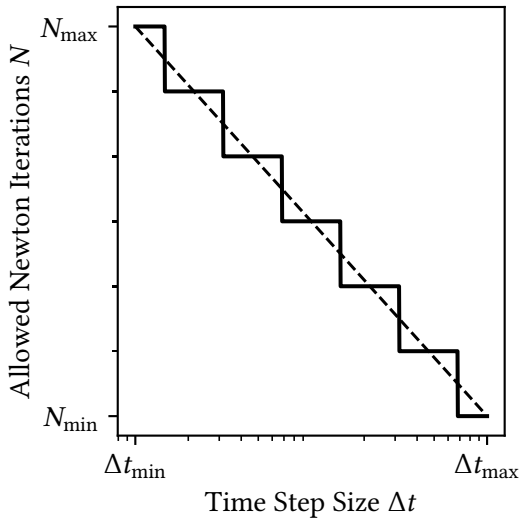

Figure 2.1: Number of allowed Newton solver iterations for a given time step size for an exemplary setting. Note the logarithmic x-axis. The parameters of the function defining the iterations (dashed line) are specified by the user. Its values are rounded to yield the allowed number of Newton solver iterations N (solid line) for a given time step size Δt.

therefore allow only a certain number of Newton iterations per stage of the time step scheme based on the logarithm of the time step size (Figure 2.1). If the solver does not converge, the solution is discarded, the time step is reduced, and the solver is applied again. If the solver converges, the time step is increased. This leads to a small time step and many allowed solver iterations for strong dynamics and long time steps with few solver iterations for weak dynamics, and allows for covering the wide temporal range of phenomena observed in the Richards regime.

There are several ways of improving this method. Discarding solutions and re-computing them for smaller time steps is costly and inefficient. A more elaborate method could accept the solution with a lower precision under certain conditions and subsequently reduce the time step. Additionally, time step sizes could be based on a prior time step estimation, similar to the pore-scale equilibration timescale discussed by Roth (2008). Finally, time step size could be based on an error measure of the resulting solution, independent of the performance of the Newton method.

2.7 Flux Reconstruction

Notably, the solution u_h of both the DG and the FV discretization cannot be considered a matric head h_m in the physical sense. This is most obvious for FV, where the potential on each grid cell is specified by only a single value. Here, a grid cell T can be interpreted as averaging volume, with its respective value u_T being a local average of the higher-order "true" function u^*, $u_T = \int_T u^* \, \mathrm{d}x$. Recalling the considerations of the REV scale in section 1.2, this requires any grid cell to be smaller than the MAV. In a DG function space, the effective

resolution additionally depends on the polynomial order of the function space. Grid cells may therefore be generally larger than the MAV, but there exists no clear limit as to how much. In practice, we find that the grid resolution must be high enough to adequately represent occurring infiltration fronts for a correct estimation of the local conductivity.

It is less obvious that the solution of the DG formulation also does not conform to a "physical" solution. While the scheme penalizes, and thus tries to minimize, the solution jump $[[u_h]]$ at grid faces, small-scale jumps typically remain. In extreme cases, like intense Dirichlet boundary conditions, these jumps can even be macroscopic, in the sense that their magnitude is similar to that of the solution value. To recover a continuous solution, potential reconstruction can be applied, which computes the average of potentials at each grid or interpolation node (Di Pietro and Ern 2012).

Flux reconstruction computes a flux with continuous normal components at grid interfaces by projecting the solution onto a Raviart-Thomas function space. The flux $\boldsymbol{j}_h$ computed directly from the solution u_h of the weak problem,

$$\boldsymbol{j}_h(u_h) = -K(u_h)\left[\nabla u_h - \hat{\boldsymbol{g}}\right], \tag{2.32}$$

is not conforming because the jump of its normal components across grid interfaces is non-zero,

$$[[\boldsymbol{j}_h(u_h) \cdot \hat{\boldsymbol{n}}_F]] \neq 0. \tag{2.33}$$

The $H(\mathrm{div})$-conforming flux in Ω can be computed from the solution u_h by prescription of a vector-valued Raviart-Thomas function space defined by

$$\mathbb{RT}_h^k = \left\{\boldsymbol{v}_h \in H(\mathrm{div};\Omega) : \forall T \in \mathcal{T}_h, \boldsymbol{v}_h|_T \in \mathbb{RT}_k(T)\right\}, \tag{2.34}$$

with $\mathbb{RT}_k(T) = [\mathcal{P}_k(T)]^d + \boldsymbol{x}\mathcal{P}_k(T)$.

For this flux prescription, we first define the numerical fluxes $\phi_h(u_h)$ at grid faces $F \in \mathcal{F}_h$ for both discretization schemes. In the FV scheme, the numerical flux is given by the contributions at each face,

$$\phi_h^{\mathrm{FV}} \coloneqq \begin{cases} \langle K_*^+, K_*^- \rangle \left[[[u_h]]/\ell_{T^\pm} - \hat{\boldsymbol{g}} \cdot \hat{\boldsymbol{n}}_F\right] & \text{if } F \in \mathcal{F}_h^i, \\ K_*^+ \left[[[u_h]]/\ell_{T,F} - \hat{\boldsymbol{g}} \cdot \hat{\boldsymbol{n}}_F\right] & \text{if } F \in \mathcal{F}_h^D, \\ j_N & \text{if } F \in \mathcal{F}_h^N. \end{cases} \tag{2.35}$$

The numerical flux in the DG scheme consists of the contributions of the consistency term, eq. (2.14b), and of the penalty term, eq. (2.14d), and the equivalent contributions from the

boundary conditions (Di Pietro and Ern 2012),

$$\phi_h^{\mathrm{DG}} \coloneqq \begin{cases} -\{\!\{K\left[\nabla u_h - \hat{\boldsymbol{g}}\right]\}\!\}_{\omega_K} \cdot \hat{\boldsymbol{n}}_F + \eta \langle K^-, K^+ \rangle [\![u_h]\!] / h_F & \text{if } F \in \mathcal{F}_h^i, \\ -K_*^+ \left[\nabla u_h - \hat{\boldsymbol{g}}\right] \cdot \hat{\boldsymbol{n}}_F + \eta K_*^+ [\![u_h]\!] / h_F & \text{if } F \in \mathcal{F}_h^D, \\ j_N & \text{if } F \in \mathcal{F}_h^N. \end{cases} \tag{2.36}$$

Following Ern, Nicaise, and Vohralík (2007), the reconstructed flux $\boldsymbol{r}_h \in \mathbb{RT}_h^k$ is prescribed with

$$(\boldsymbol{t}_h \cdot \hat{\boldsymbol{n}}_F, q_h)_F = (q_h, \phi_h^{\mathrm{m}}(u_h)), \quad \mathrm{m} \in \{\mathrm{FV}, \mathrm{DG}\} \tag{2.37}$$

for all $F \in \mathcal{F}_h, q_h \in \mathcal{P}_k(F)$, and in case of the DG scheme additionally with

$$\begin{aligned} (\boldsymbol{t}_h, \boldsymbol{r}_h)_T = -(\boldsymbol{r}_h, K\left[\nabla u_h - \hat{\boldsymbol{g}}\right])_T &+ \sigma \sum_{F \in \partial T \cap \mathcal{F}_h^i} \left([\![u_h]\!], \left\langle K^-, K^+ \right\rangle \boldsymbol{r}_h \cdot \hat{\boldsymbol{n}}_F \right)_F \\ &+ \sigma \sum_{F \in \partial T \cap \mathcal{F}_h^D} \left([\![u_h]\!], K_*^+ \boldsymbol{r}_h \cdot \hat{\boldsymbol{n}}_F \right)_F \end{aligned} \tag{2.38}$$

for all $T \in \mathcal{T}_h, \boldsymbol{r}_h \in [\mathcal{P}_{k-1}(T)]^d$. When considering numerical solutions in the following, I will denote fluxes calculated from the solution gradients via eq. (2.32) with $\boldsymbol{j}_h$ and reconstructed fluxes with $\boldsymbol{j}$.

We find that the prescription terms coincide directly with terms of the FV and DG space semi-discretization forms given by eqs. (2.11) and (2.14), respectively, albeit with inverted signs in the volume contributions and a different definition of the test functions. Therefore, any operator written for applying these forms can be used for prescribing the reconstructed flux with only few modifications (Ospina De Los Ríos 2019). However, the given formulation is only applicable to matching meshes and therefore cannot handle hanging nodes on the grid. Flux reconstruction on non-matching grids can be performed by solving a minimization problem formulated with local Neumann problems in a mixed finite element setting (Ern and Vohralík 2009). The approach of representing fluxes across grid interfaces with vector-valued function spaces is also used in hybrid DG methods, which further increase the efficiency of the method by minimizing the number of globally coupled DOFs compared to regular DG methods (Cockburn, Gopalakrishnan, and Lazarov 2009).

2.8 Adaptive Grid Refinement

Applying adaptive grid refinement, or h-refinement, for solving the Richards equation has proven to increase accuracy, efficiency, and even robustness of the numerical method (Li, Farthing, and Miller 2007). The efficiency of DG methods additionally strongly profits from simultaneous p-refinement, where the polynomial order of the local function space on

single grid elements is adapted according to an error measure. Solin and Kuraz (2011) apply hp-refinement on the Richards equation and find that it improves the spatial accuracy of the solution, but also that h-refinement with intermediate-order polynomial basis functions is competitive if high spatial accuracy is not required. Combined hp-refinement has successfully been applied with a DG method for two-phase flow in DUNE (Dedner et al. 2018). Nonetheless, p-refinement requires a highly flexible, specialized implementation, which is necessarily less optimized in terms of memory layout and access. Simpler numerical discretizations like FV methods typically do not even support computations on the resulting irregular grids. With DG methods still being rarely used for soil water flow problems, h- and p-adaptivity are uncommon in practical implementations (Miller et al. 2013).

Error indicators for grid refinement usually operate on the resulting solution and incorporate knowledge of the discretization scheme. For the SWIP DG method, Ern and Stephansen (2008) derive an error upper bound of the solution based on three indicators, (i) the residual error based on a projection of the residual onto piecewise constant function space, (ii) the nonconforming error based on the difference between the solution and the reconstructed solution, and (iii) the diffusive flux error based on the interface jumps of the solution and the solution gradient. The latter is the simplest to evaluate because it requires no reconstruction or projection and is localized to single grid elements T. Following the formulation of Di Pietro and Ern (2012), we use it as local error indicator,

$$\mathcal{E}_{\mathrm{df},T} \coloneqq \sqrt{C_{F,T} h_T} \left\| \frac{1}{2} [\![K \nabla u_h]\!] \cdot \hat{\boldsymbol{n}}_T + \frac{\eta}{h_F} \left\langle K^+, K^- \right\rangle [\![u_h]\!] \right\|_{\partial T}, \tag{2.39}$$

with

$$C_{F,T} = \left[h_T |\partial T| |T|^{-1} \right] \left[2d^{-1} + \pi^{-1} \right] \pi^{-1},$$

where $\|\cdot\|_{\partial T}$ is the L^2-norm on the grid cell boundary ∂T, and $[\![K \nabla u_h]\!] = 0$ on boundary faces $F \in \mathcal{F}_h^b$. The indicator has the unit of a flux, $[\mathcal{E}_{\mathrm{df},T}] = \mathrm{m\,s^{-1}}$, and can be interpreted as erroneous numerical flux due to non-physical jumps in the solution and its gradient.

With two error bounds $\alpha_T > \beta_T$, a threshold algorithm for grid refinement will split a grid element T if $\mathcal{E}_{\mathrm{df},T} > \alpha_T$, and recombine two adjacent, refined grid elements $T^{\pm}$ with common "ancestor" if $\mathcal{E}_{\mathrm{df},T^{\pm}} < \beta_T$. Then the solution is projected onto the new grid. If the grid was refined, the typical procedure is to recompute the solution for the current time step on the new grid, and to reiterate this procedure until all local errors are below the refinement threshold α_T, or a global error limit is reached (Li, Farthing, and Miller 2007). This procedure ensures a consistent, maximum error bound for all time steps, but is computationally expensive.

We only apply a single adaptation step after each time step, and then continue the computation for the next time step using the new grid. This means that the grid adaptation generally lags the dynamics, which has to be counteracted by a sufficiently small refinement

threshold α_T and adequately small time steps. Additionally, the projection of the solution onto a coarsened grid preserves the projected quantity $u_h \equiv h_m$, but not the water mass represented by $\theta(h_m)$. The error resulting from this violation of mass conservation can be minimized by choosing a sufficiently small coarsening threshold β_T, which ensures that the local state is near equilibrium. On the other hand, computational costs may increase greatly if the thresholds are chosen too low and local refinement of the grid is unnecessarily strong.

3 DORiE

3.1 Overview

> This section is based on, and extends, Riedel et al. (2020a): "DORiE: A discontinuous Galerkin solver for soil water flow and passive solute transport based on DUNE."

DORiE is a module of the DUNE framework (Blatt et al. 2016; Bastian et al. 2020) and its high-level discretization module DUNE-PDELab (Bastian, Heimann, and Marnach 2010). It implements the numerics of chapter 2 on structured rectangular and unstructured rectangular or simplex grids in two and three spatial dimensions. The FV solver always uses the structured rectangular YASP grid manager supplied by the DUNE-Grid module. For the DG solver, DORiE additionally employs the UG grid manager[1] for unstructured grids. While rectangular grids are always created within the application based on specifications by the user, unstructured simplex grids are loaded from GMSH mesh files (Geuzaine and Remacle 2009). If local grid refinement is enabled, DORiE must use the unstructured grid manager that causes a computational overhead compared to using the structured grid manager on the same grid. DORiE currently strictly separates grids with simplex and rectangular shapes. Therefore, h-refinement on rectangular grids does not use closures and leads to hanging nodes on the grid, which invalidates the reconstructed flux. For every combination of numerical discretization, grid type, and polynomial order (in case of DG), DORiE compiles an optimized program instance. This increases compile time but avoids any overhead when executing the program with a certain setting.

DORiE provides a Python command line interface for user interaction. The main settings for a simulation are supplied through an INI[2] file in a format specification by the DUNE framework. Therein, users supply the grid specification, discretization scheme settings, solver precision, time stepping parameters, and data file paths, among others. Parameterization and boundary condition data files are supplied in the hierarchical YAML[3] format. DORiE currently supports the Mualem–van Genuchten parameterization and a linear interpolation of boundary condition values in time. Additional data like initial conditions, scaling

1. See https://gitlab.dune-project.org/staging/dune-uggrid.
2. Neither an abbreviation, nor a standardized format. See https://en.wikipedia.org/wiki/INI_file.
3. See https://yaml.org/.

fields, and domain mappings can be specified in HDF5 files (The HDF Group 1997–2020). Initial conditions can also be computed from the stationary problem through an initial set of boundary conditions. Simulation output is written in the Visualization Toolkit (VTK; Schroeder, Martin, and Lorensen 2006) file format and can be analyzed and visualized with third-party software or the Python package included in DORiE.

For a convenient definition of heterogeneity and soil structure, DORiE uses domain mappings. Every grid cell $T \in \mathcal{T}_h$ receives and index $t \in \mathbb{Z}$, and every boundary face $F \in \mathcal{F}_h^b$ an index $f \in \mathbb{Z}$. These indices are used to identify the parameterization of the grid cell and the boundary condition of the face, respectively. For rectangular grids, users supply these indices as single $\mathbb{Z}^d$ dataset for cell mappings and $2d \times \mathbb{Z}^{d-1}$ datasets for boundaries in HDF5 files. GMSH supports a mapping of "physical" entities natively and saves these mappings into its mesh files, which means that no additional information is required when importing a simplex mesh. Boundary conditions and parameterizations are then assigned by repeating the respective indices in the YAML data files. If a scaling field is supplied, the (interpolated) field value at the barycenter of any grid cell is stored as scaling factor for this grid cell. Scaling factors and mapping indices t and f are inherited by "child" grid elements under local grid refinement.

DORiE also features a solver for passive solute transport, which has been coupled with the Richards solver module and can be enabled optionally (Ospina De Los Ríos 2019). Additionally, it integrates the Gaussian random field generator based on circulant embedding of the DUNE-Randomfield module (Klein 2016). The code, documentation, and functionality of DORiE version 2.0 have been reviewed as part of the publication in the Journal of Open Source Software (see Riedel et al. 2020a).

3.2 Benchmarks

I investigated the accuracy and efficiency of the numerical schemes implemented in DORiE with three different scenarios. First, I solved the stationary, pseudo one-dimensional problem and compared it to a highly resolved reference solution. I then studied the resolution of transient features with the example of an infiltration front passing through a layered medium. Finally, I analyzed the efficiency of the transient solver with a benchmark on a heterogeneous domain, including varying boundary conditions and h-refinement.

3.2.1 Solution Convergence

I surveyed the convergence of the solution towards a highly resolved reference for an increasing grid resolution, i.e., a decreasing grid cell extension h_T. Given an exact solution $u \in H^{k+1}$ that solves eq. (2.6), where H^{k+1} is the Sobolev space of functions with derivatives

up to order $k+1$, we expect a proportional relationship between the L^2-error of the numerical solution $u_h \in V_h$ and the cell extension to the power $k+1$ for the DG scheme (Ern, Stephansen, and Zunino 2009),

$$\|u_h - u\| \propto h_T^{k+1}, \tag{3.1}$$

where the L^2-error of a function on the physical domain $\Omega \subset \mathbb{R}^d$ is defined as

$$\|\cdot\| \coloneqq \sqrt{\int_\Omega |\cdot|^2 \,\mathrm{d}x}. \tag{3.2}$$

As no general analytical solutions of the Richards equation with Mualem–van Genuchten parameterization exist, we have to estimate the true solution u. In a stationary system, where the time derivatives are zero, the local flux is given by the Buckingham-Darcy law. Assuming a horizontally homogeneous system, and that gravity acts in the vertical direction, the vertical flux is then given by

$$j_z = -K(h_m)\left[\partial_z h_m - 1\right], \tag{3.3}$$

where the subscript z indicates the vertical axis. By convention, the z encodes the depth and the associated axis points downwards, with the soil surface at depth $z = 0\,\mathrm{m}$. In a stationary system, the flux j_z is constant over depth, $\partial_z j_z = 0$. The presented equation thus is an ODE for the matric head in the vertical direction,

$$\mathrm{d}_z h_m = 1 - \frac{j_0}{K(h_m)}. \tag{3.4}$$

This equation can only be solved analytically if the expression for the conductivity $K(h_m)$ is simplified. However, it can be solved numerically with a simple ODE integration algorithm given the potential at the initial (largest) depth z_0, $h_m(z = z_0)$, and the flux through the domain j_0. As the numerical implementation for ODE solvers is far simpler than the FV and DG discretization schemes, I assumed it to yield a more accurate solution for a stationary, one-dimensional problem.

I tested the convergence by computing a stationary solution on a pseudo one-dimensional grid with one grid cell in horizontal direction and 5×2^n cells in the vertical direction, with $n = 0, \ldots, 8$. The grid spanned a domain of $1\,\mathrm{m} \times 1\,\mathrm{m}$. At the bottom, I used a Dirichlet boundary condition to set a constant water table, $h_m(z = 1\,\mathrm{m}) = 0\,\mathrm{m}$. This also served as initial condition value for the ODE integrator. At the top, I set a Neumann boundary condition with the desired flux for the entire domain, $j_N(z = 0\,\mathrm{m}) = -5 \times 10^{-8}\,\mathrm{m\,s^{-1}}$, which I also inserted into the ODE integrator as flux throughout the domain via $j_0 = -j_N$. The vertical boundaries had no-flow boundary conditions with $j_N = 0\,\mathrm{m\,s^{-1}}$. I expected the parameterization functions to have a profound influence on the solver accuracy, so I ran the simulations for four different sets of parameters, namely *Loam*, *Sandy Loam*, *Loamy Sand*,

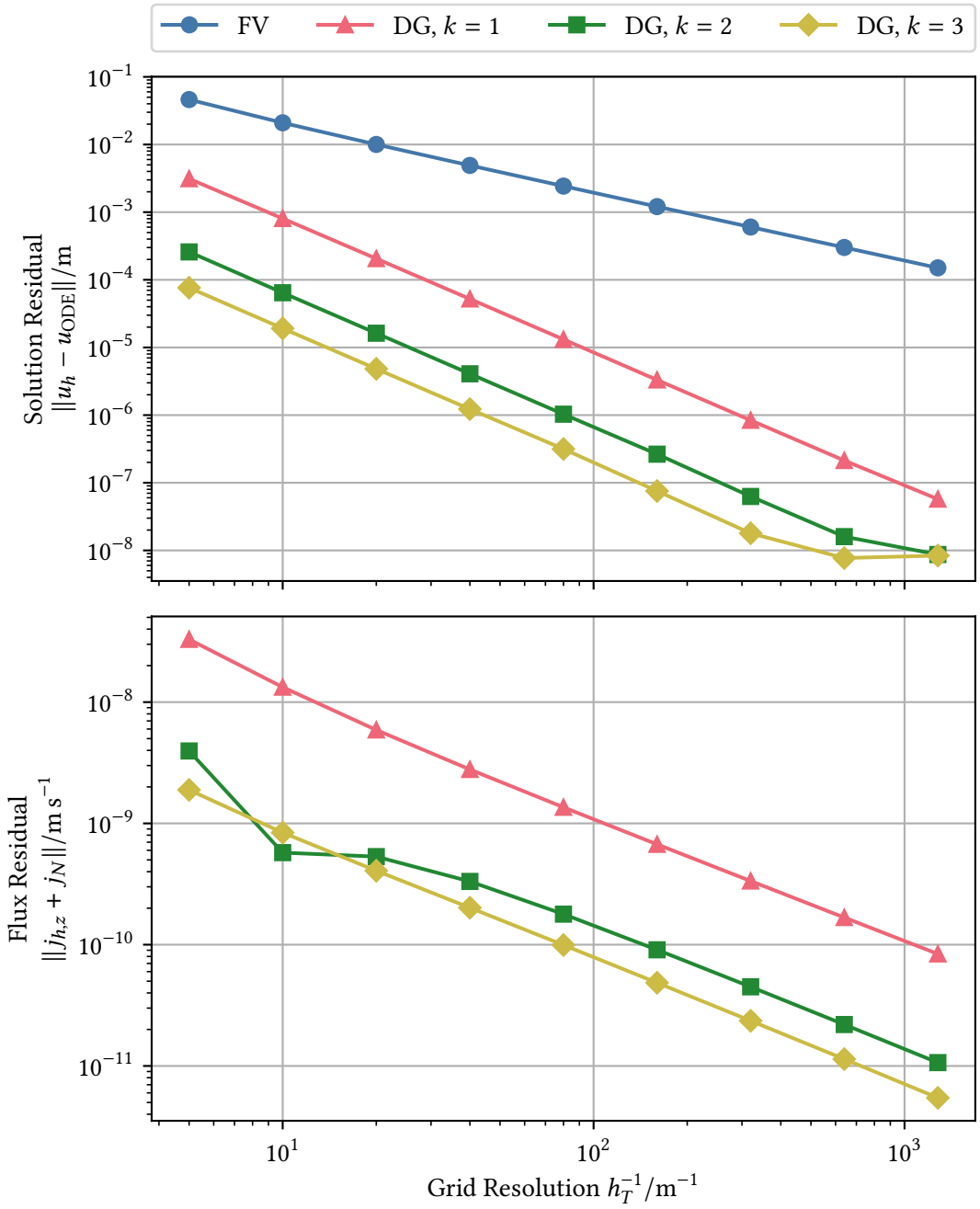

Figure 3.1: Exemplary convergence plots of the solution residual (top) and the flux residual (bottom) against the grid resolution for the FV and DG schemes, and, in case of DG, for three different polynomial orders k. The solution residual is computed by the difference to the numerical ODE solver solution u_{ODE} and the flux residual by the difference between the vertical solution flux $j_{h,z}$ and the negative boundary condition flux $-j_N$. The parameter set of this case is *Loam*.

Table 3.1: Solution residual convergence rates of FV and DG in three polynomial orders k for four parameter sets.

Medium	Solution Residual Convergence Rate			
	FV	DG, $k = 1$	DG, $k = 2$	DG, $k = 3$
Loam	1.03 ± 0.01	1.97 ± 0.00	1.92 ± 0.04	1.78 ± 0.10
Loamy Sand	1.04 ± 0.01	1.96 ± 0.02	2.11 ± 0.07	1.98 ± 0.07
Sandy Loam	1.03 ± 0.01	2.01 ± 0.00	1.91 ± 0.03	1.81 ± 0.07
Silt Loam	1.02 ± 0.01	1.95 ± 0.00	1.82 ± 0.04	1.55 ± 0.12
Mean	1.03 ± 0.01	1.97 ± 0.02	1.94 ± 0.12	1.78 ± 0.18

Table 3.2: Flux residual convergence rates of DG in three polynomial orders k for four parameter sets.

Medium	Flux Residual Convergence Rate		
	DG, $k = 1$	DG, $k = 2$	DG, $k = 3$
Loam	1.06 ± 0.02	0.95 ± 0.07	1.04 ± 0.01
Loamy Sand	1.05 ± 0.03	2.09 ± 0.05	2.39 ± 0.16
Sandy Loam	1.07 ± 0.03	1.93 ± 0.08	1.80 ± 0.12
Silt Loam	0.98 ± 0.01	0.82 ± 0.03	0.89 ± 0.00

and *Silt Loam*, cf. table 1.1. The particular value of the Neumann boundary condition was chosen to pose no numerical difficulties for the particularly stiff stationary problem in any of the scenarios. Further investigations revealed that the value of the boundary condition influences the absolute value of the solution residual, but not the residual convergence rate.

The solution residual was computed with the L^2-norm of the difference between the DORiE solution u_h and the ODE integrator solution u_{ODE}. I chose the fourth-/fifth-order Dormand-Prince Runge-Kutta (DOPRI5)[4] method as ODE integrator, which is readily implemented in the Python SciPy package (Virtanen et al. 2020). This method features an automated integration step size control based on an absolute tolerance of 10^{-12} and a relative tolerance of 10^{-6} for the error between the results of the fourth- and fifth-order integration. For the flux residual, I calculated the L^2-norm of the difference between the vertical solution flux $j_{h,z}$ and the negative boundary condition flux $-j_N$. The solution flux $j_{h,z}$ was calculated from the solution u_h according to eq. (2.32) and directly represents the numerical fluxes, cf. section 2.7. The residual of the reconstructed flux was always below the numerical precision

4. For the documentation and source code, see https://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.ode.html.

limit in this scenario, so I did not investigate it further. For FV solutions, local gradients ∇u_h are always zero, and the solution flux cannot be calculated without flux reconstruction.

The mean solution convergence rate is 1.03 ± 0.02 for FV and 1.97 ± 0.02 for DG with polynomial order $k = 1$ (Table 3.1). This coincides with the expected rates, cf. eq. (3.1). DG with orders $k > 1$ do not achieve a higher convergence rate but only an offset of the solution residual, which decreases with increasing polynomial order (Figure 3.1). I attribute this behavior to the reduced regularity induced by the conductivity function. Convergence rates even partially decrease for higher polynomial orders, most notably for *Silt Loam*. For high grid resolutions and polynomial orders, the solution residual reaches an asymptotic limit. This is due to the single-precision VTK output provided by the DUNE backend: Since the solution is in the order of $O(u_h) \approx 1\,\mathrm{m}$, the numerical precision limit is expected to be about $10^{-8}\,\mathrm{m}$. Still, higher-order methods always have a lower residual than lower-order methods at the same grid resolution, and increasing the grid resolution always increases the accuracy.

Flux convergence rates are usually higher than 1, but do not follow a clear pattern (Table 3.2). As the flux residual is based on the gradient of the solution, it is more sensitive to oscillations, which are a typical feature of numerical solutions. Especially in the low resolution regime, the flux residual therefore behaves unexpectedly. For *Loam*, the flux residual for $k = 2$ was lower than for $k = 3$ at certain low grid resolutions (Figure 3.1). At this particular instance the setup apparently lead to comparably low oscillations for the lower order. Nonetheless, the flux residuals always decreased with increasing grid resolution and overall showed a smooth convergence, as indicated by the mostly low uncertainty of the fitted convergence rates displayed in table 3.2.

3.2.2 Infiltration Front Resolution

I investigated the resolution of a sharp infiltration front by employing the instationary solver on a pseudo one-dimensional setup similar to the one used in the last section. The regular grid covered a $1\,\mathrm{m} \times 1\,\mathrm{m}$ domain with one grid cell in horizontal direction and 40×2^n cells in vertical direction with $n = 0, \dots 4$. I denoted the vertical extension of the grid cells in the vertical direction with h_T. The domain contained two media and an interface at depth $z = 50\,\mathrm{cm}$; the upper medium was homogeneous *Sandy Loam*, the lower medium was homogeneous *Loamy Sand*. The initial condition throughout the domain was gravity flow, $\partial_z h_m = 0$, with a constant potential of $h_m = -1\,\mathrm{m}$. The same value was chosen as lower boundary condition. The upper boundary condition was a constant Neumann infiltration flux of $j_N = -10^{-6}\,\mathrm{m\,s^{-1}} = -86.4\,\mathrm{mm\,d^{-1}}$. To achieve results independent of the temporal integration, I chose a comparably small, fixed time step size $\Delta t = 200\,\mathrm{s}$ for all simulations. I applied the DG scheme with $k = 1, 2, 3$ and the FV scheme. Since I established in the previous section that DORiE produces consistent results and its accuracy increases with

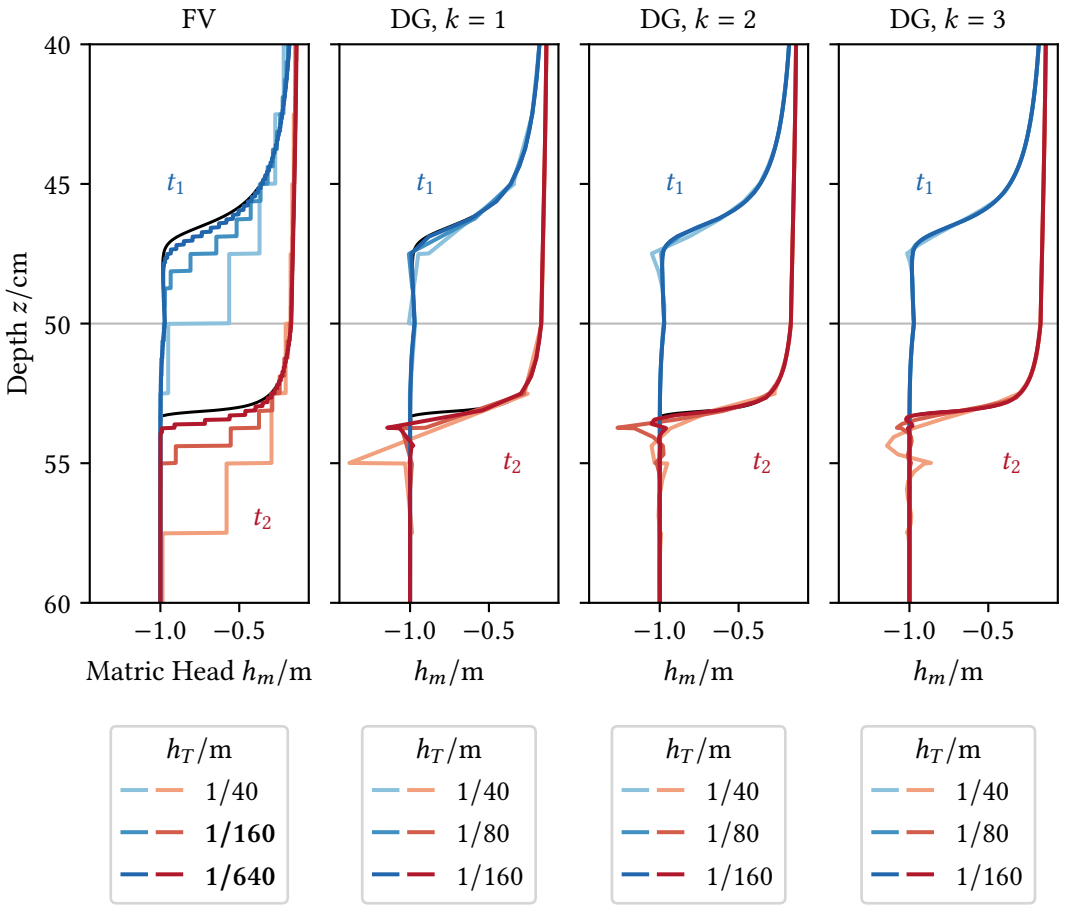

Figure 3.2: Water front infiltrating into a layered medium, depicted at two times $t_1 = 8.5 \times 10^4$ s and $t_2 = 10^5$ s for different grid resolutions and discretization methods. The layer boundary between *Sandy Loam* (top) and *Loamy Sand* (bottom) is indicated by the gray line at $z = 50$ cm. Main colors denote the two different points in time (blue: t_1; red: t_2) and color shades denote three different grid resolutions h_T, as indicated in the legends. Note that the grid resolutions vary for FV and DG results; bold entries for FV indicate deviating values. A high-resolution result from a DG simulation with $k = 3$, $h_T = 1/640$ m is plotted as two thin black lines (partially obscured) in every panel for comparison.

increasing grid resolution and polynomial order k, I used the most accurate solution with $k = 3$, $h_T = 1/640\,\mathrm{m}$ as reference for the results with lower resolution.

The matric head over depth from the results of all discretization schemes and three grid resolutions each for two points in time, $t_1 = 8.5 \times 10^4\,\mathrm{s}$ and $t_2 = 10^5\,\mathrm{s}$, are visualized in fig. 3.2. At $t_1 \approx 23.3\,\mathrm{h}$, the water front is about to reach the medium interface at $z = 50\,\mathrm{cm}$. We find that the velocity of the water front is more overestimated with lower grid resolutions and lower-order methods. This can be seen most clearly for FV, $h_T = 1/40\,\mathrm{m}$, where the matric head below the layer interface already begins to rise. The infiltration depth of the water front is overestimated, and the potential at the trailing side of the front is underestimated, which is the equivalent of an overestimated conductivity $K(h_m)$, leading to a lower rise in water content $\Delta\theta$ and a consequently faster front propagation speed $v \approx |j_N|\Delta\theta$. The effect decreases with increasing grid resolution.

The DG method is again more accurate than the FV method. Notably, $k = 1$ at $h_T = 1/160\,\mathrm{m}$ achieves a higher accuracy than FV at $h_T = 1/640\,\mathrm{m}$ (the quadruple grid resolution), which is also supported by the results of the stationary case, cf. fig. 3.1. Although the infiltration front propagation speed is also overestimated for low grid resolutions, this effect is less pronounced than for FV and reduces with increasing order k. However, especially for low resolutions, the DG solution features over- and undershoots not present in the FV solution, which become even more pronounced after the front has passed the layer interface and entered the *Loamy Sand* at $t_2 \approx 27.8\,\mathrm{h}$.

These features are a result of the DG method and its high number of DOFs per grid element. (Numerical) fluxes between the grid elements are given by solution jumps $[[u_h]]$ and solution gradients ∇u_h at the interface, with the mass balance directly influencing the local water content and hydraulic conductivity. Propagating water into a grid cell with low water content and hence low conductivity requires a high gradient in the matric head, but the gradient and the jumps of the solution at interfaces in turn limit the overall water content of the grid cell. This leads to discrepancies if a sharp water front enters a grid cell with comparably large extensions. The required strong gradient overshoots, and the solution jump to the next cell produces a small reverse flux at the other side of the cell, see fig. 3.2, DG, $k = 1$, $t = 8.5 \times 10^4\,\mathrm{s}$, $h_T = 1/40\,\mathrm{m}$. As long as this non-physical flux is not too strong, it stabilizes the method because it increases the water content in the cell with water from below the infiltration front. While the front propagates further, the overshooting potential then "unravels" and increases until the infiltration flux can be supported without a local gradient, which is the consistent solution for the gravity flow regime, cf. section 1.6.

Overshoots are a typical feature of convective high-order methods. They can be avoided with slope limiters, which apply a projection on the residual in each stage of the Runge-Kutta method (Burbeau, Sagaut, and Bruneau 2001). This could be a useful addition to a future version of DORiE.

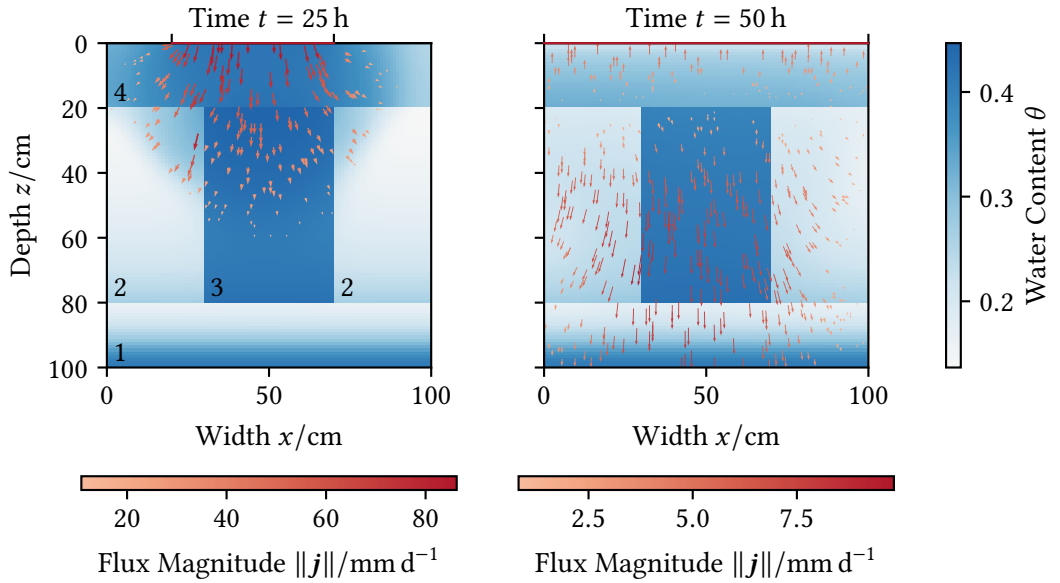

Figure 3.3: Reference solution of the benchmark setup right before the infiltration stops (left), and at the end of the simulation (right). The water content color bar applies to both panels. Jumps in the water content denote material interfaces, with the numbers in the left panel indicating the media *Loamy Sand* (1), *Sandy Loam* (2), *Silt Loam* (3), and *Loam* (4), cf. table 1.1. The surface segment at which the respective boundary condition applies is indicated by the red line at the top axis.

3.2.3 Transient Setting with Grid Refinement

I analyzed the accuracy and performance of DORiE using transient simulations with a two-dimensional, heterogeneous domain. This 1 m × 1 m domain was divided into three layers with a respective depth of 20 cm, 60 cm, and 20 cm. The top layer consisted of *Loam* and the bottom layer of *Loamy Sand*. The center layer was horizontally divided at widths of 30 cm and 70 cm, with the outer layers consisting of *Sandy Loam* and the inner layer of *Silt Loam*, cf. table 1.1 for the parameter sets. All layers were homogeneous. See fig. 3.3 for the domain division. The lower boundary condition was a constant water table, $h_m(z = 1\,\mathrm{m}) = 0\,\mathrm{m}$. At the upper boundary, I applied a constant strong infiltration of 86.4 mm d^{-1} for 25 h, followed by an outflow boundary condition which increased linearly from $h_m = 0\,\mathrm{m}$ to a value of $-5\,\mathrm{m}$ over the following 25 h. The infiltration boundary condition was only applied at a width from 20 cm to 70 cm, whereas the outflow boundary condition was applied over the entire upper boundary. I set no-flow boundary conditions at the lateral borders of the domain. The initial condition was hydrostatic equilibrium with the water table at the lower boundary.

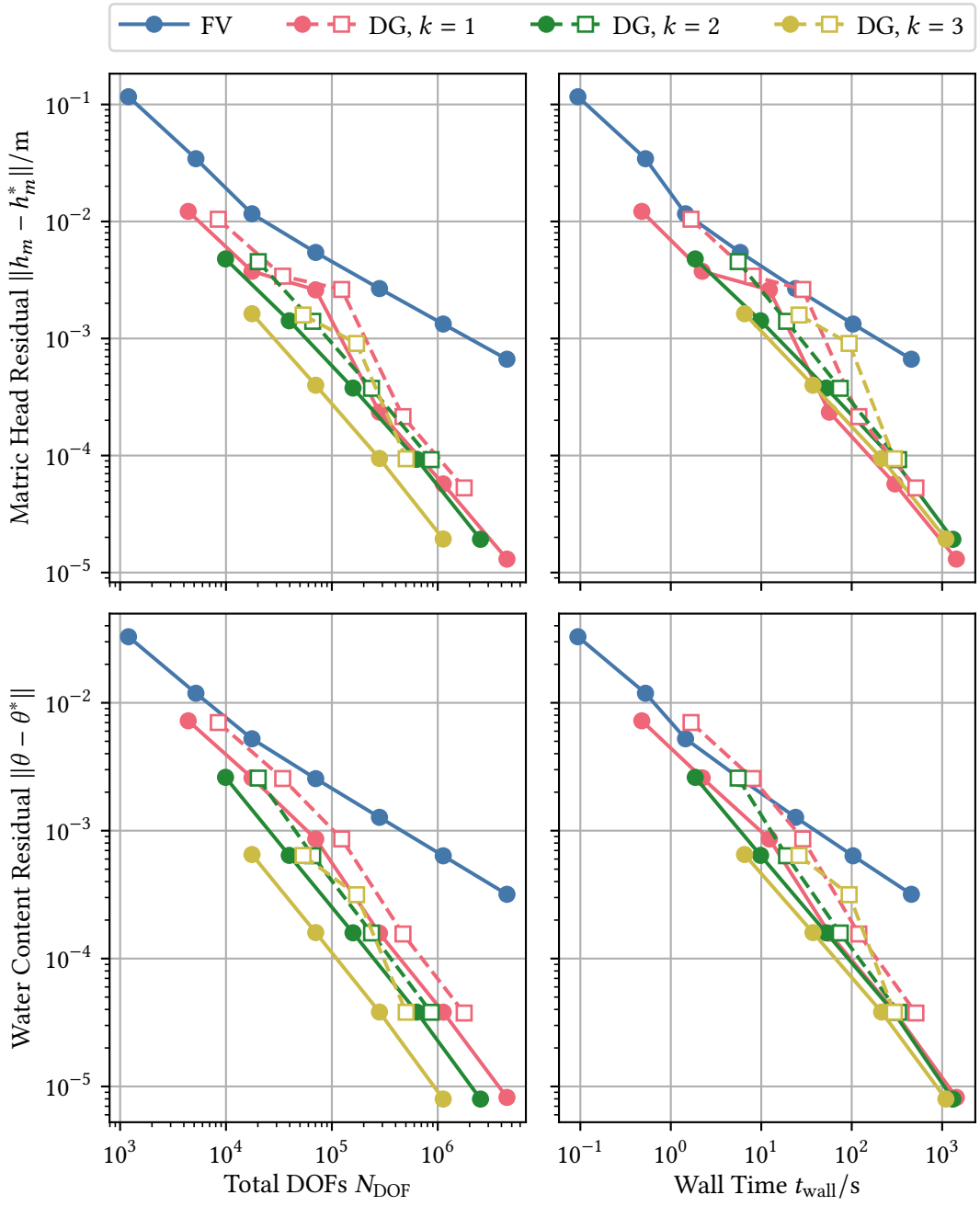

Figure 3.4: Results of the transient benchmark for simulation time $t = 25\,\mathrm{h}$, right before infiltration stops. Filled markers connected by solid lines denote results from structured, rectangular grids and outlined markers connected by dashed lines denote results from adaptive, rectangular grids. The total number of DOFs N_{DOF} was calculated by summing over all DOFs computed in $N_{t=25\,\mathrm{h}}$ time steps.

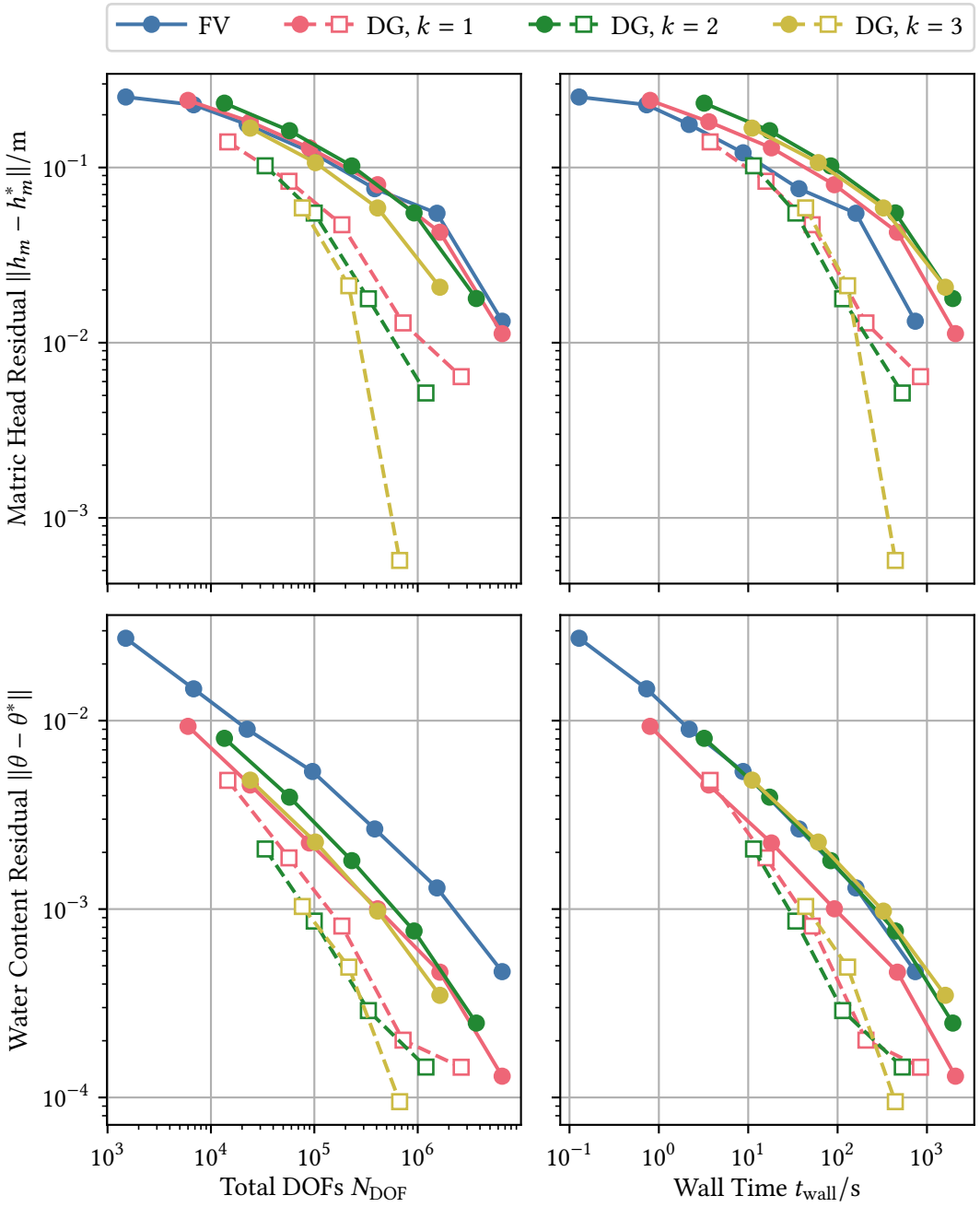

Figure 3.5: Results of the transient benchmark for simulation time $t = 50\,\mathrm{h}$, after 25 h of evaporation. Filled markers connected by solid lines denote results from structured, rectangular grids and outlined markers connected by dashed lines denote results from adaptive, rectangular grids. The total number of DOFs N_{DOF} was calculated by summing over all DOFs computed in $N_{t=50\,\mathrm{h}}$ time steps.

I chose this artificial setup for several reasons: The fine-grained *Loam* at the top allows a comparably strong evaporation flux. The more coarse-grained materials of the center layer produce sharp water fronts for strong infiltration fluxes, and feature different water contents and front propagation speeds. The coarse-grained *Loamy Sand* in the bottom layer achieves a decoupling from the water table at the bottom boundary, because its water content drops relatively quickly with decreasing potential. The application of the infiltration boundary condition additionally breaks the symmetry of the setup. I expected this setup to be highly sensitive to the grid resolution, because strong gradients would appear at the tip of the infiltration front and at the upper boundary during evaporation. Additionally, different time step sizes could manifest in errors in the front propagation and the evaporation flux, with the latter effect being amplified by the time dependency of the boundary condition.

I simulated a total of 50 h and selected the heuristic time step adjustment with a minimum time step of 1 s, a starting time step of 100 s and a maximum time step of 10^5 s (approx. 27.8 h). I ran simulations with varying grid resolutions and discretization schemes, and compared solutions from structured rectangular grids and from adaptive, unstructured grids with a reference solution. Structured grids used $(10 \times 10) \times 4^n$ cells with $n = 0, \dots, 6$ for the FV discretization, $n = 0, \dots, 5$ for the DG discretization with order $k = 1$, and $n = 0, \dots, 4$ for orders $k = 2, 3$. The solution (h_m^*, θ^*) of the combination $(k = 3, n = 4)$ was used as reference for computing the residuals. Adaptive grids started with a global refinement of level n, and allowed for a minimum local refinement level of $n - 1$ and a maximum local refinement level of $n + 1$. The adaptivity algorithm used was a threshold algorithm as explained in section 2.8, with refinement threshold α_T ranging from $10^{-8}\,\mathrm{m\,s^{-1}}$ to $10^{-10}\,\mathrm{m\,s^{-1}}$ and coarsening threshold β_T ranging from $5 \times 10^{-10}\,\mathrm{m\,s^{-1}}$ to $5 \times 10^{-12}\,\mathrm{m\,s^{-1}}$, where the particular values selected for a simulation were defined by the order k and the initial grid level n.

The L^2-residuals of the matric head h_m and the water content θ were evaluated right before the infiltration stops at $t = 25\,\mathrm{h}$, and at $t = 50\,\mathrm{h}$, the end time of the simulation. The reference solutions at these times are visualized in fig. 3.3. To indicate the performance, I compared the residuals against the simulation wall time t_{wall} taken up to retrieve the respective solution, which includes the duration of grid creation, operator setup, solution output, and grid refinement (if applicable). I additionally compared the residuals against the total number of DOF computed by the simulation,

$$N_{\mathrm{DOF}} = \sum_{i}^{N_t} N_{\mathrm{DOF},i}, \tag{3.5}$$

where $N_{\mathrm{DOF},i}$ is the number of DOF computed at time step i, and N_t is the number of time steps the simulation took to compute the solution at time t.

For static grids, the residuals at time $t = 25\,\mathrm{h}$ for both matric head h_m and water content

θ show a similar behavior than those of the static benchmark in section 3.2.1 (Figure 3.4). Residuals of the DG method are lower than those of the FV methods for the same grid resolution and DOFs, and have a higher convergence rate. Residuals decrease for increasing DG orders k, but all orders used achieve similar convergence rates. DG methods also achieve a higher accuracy than the FV method when comparing simulation wall times. However, increasing the polynomial order k generally does not yield lower residuals for the same wall time.

At time $t = 50\,\mathrm{h}$, the results for static grids are less conclusive (Figure 3.5). The residuals of the matric head resemble a super-convergence, where the convergence rate increases with increasing grid resolution and number of DOFs. This can be an artifact due to the finite resolution of the reference solution. While all discretization schemes achieve similar results for the same number of DOFs when comparing residuals of the matric head, the DG schemes perform better than the FV scheme with respect to the water content. When comparing residuals based on wall time t_{wall}, the DG schemes perform similar or even worse than the FV scheme.

Adaptive grids have a worse performance than static grids when comparing solutions at time $t = 25\,\mathrm{h}$, and a better performance when comparing them at time $t = 50\,\mathrm{h}$, both in terms of number of DOFs and simulation wall time. They achieve similar or slightly better convergence rates than the method with the same respective polynomial order k on static grids. Also, residuals decrease with increasing polynomial order in most cases.

In this benchmark, the DG method only clearly outperforms the FV method on static grids when infiltration is simulated. Sufficiently strong infiltration leads to sharp infiltration fronts, which can be better resolved by the polynomial representation of the DG method. The front diffuses during evaporation, and the FV method profits from its much simpler formulation. But the results also demonstrate that the DG scheme can still be more accurate and efficient than the FV scheme in such situations if h-refinement is employed, which optimizes the number of DOFs and increases the resolution where the solution is less smooth. Nonetheless, the h-refinement algorithm shows deficits during infiltration, when the overall error mostly hinges on the resolution of the infiltration front moving through the domain. In this case, the algorithm apparently was unable to track the moving features sufficiently well, but it successfully refined upon static features of the solution.

3.3 Conclusion and Outlook

I verified the results of both discretization schemes featured in DORiE by comparing them to a highly resolved reference solution. As expected, increasing the grid resolution or the order of the discretization scheme always improves the accuracy of the solution. However, DG methods are also expected to strongly increase their accuracy with increasing polynomial

order. I could not observe this feature in our application, which is likely due to the non-regularity induced by the conductivity function, and, in the transient case, additionally by the retention curve. The properties of these parameterization functions effectively reduce the integrability and differentiability of the solution, which makes its higher-order representation less beneficial. Consequently, the computational efficiency of the DG discretization does not necessarily surpass that of the FV method. While the accuracy of the DG method is much higher than that of the FV method on the same grid, the FV method can reach an equivalent accuracy for the same number of computed DOFs, and even a better accuracy for the same wall time in certain scenarios. This deficit can be improved upon by optimizing the code base and by employing h- or p-refinement.

Adaptive refinement allows the solver to optimize the resolution during a simulation run. I demonstrate that our h-refinement method significantly outperforms computations on static grids if the solution features to be refined upon are static. In these cases, the error convergence rate for adaptive grids is higher than for static grids, which is supported by the results of Li, Farthing, and Miller (2007) and Solin and Kuraz (2011). However, our method can fail to resolve these essential features in transient scenarios with sharp infiltration fronts. Further work on DORiE therefore should improve the refinement algorithm.

Similar to Miller et al. (2013), I conclude that the DG discretization is not always advantageous to the much simpler FV method, at least not without further optimization. If the simulated domain features considerable heterogeneity and numerical accuracy is a lesser concern, the FV method is the better choice. The DG method can produce a similar accuracy in the same or lower simulation wall time on a coarser grid. But increasing the grid cell size necessarily reduces the resolution of heterogeneity in the simulated domain. The prior knowledge of heterogeneity therefore sets a lower limit to the grid resolution used during simulations. This, in turn, makes the regime of low accuracy and short simulation wall time inaccessible to the DG scheme.

Part II

Application with Data Assimilation

4 Ensemble Kalman Filter in Soil Hydrology

Forward models are used for forecasting future system states from a known state. Accurate forecasts therefore require an appropriate conceptual and numerical representation of the physical problem and an exact knowledge of the problem parameters, like boundary conditions and material properties. Due to the highly non-linear material properties encountered in soil hydrology, uncertainties in the material properties lead to significant uncertainties in the resulting forecast. As soil hydraulic material properties typically cannot be measured directly, they have to be inferred through inverse modeling. Inverse modeling is the technique of "inverting" the model and determining model parameters based on observations of the true system and simulations. Consequently, inverse modeling has been applied in soil hydrology for several decades (Vrugt et al. 2008).

The need for quantifying uncertainties in both model parameters and estimated system states led to the application of data assimilation methods in soil hydrology (Liu et al. 2012). Likewise, inverse methods have been adapted to evaluate uncertainties and representation errors (Jaumann and Roth 2017). Data assimilation and inverse modeling are considered to be strongly overlapping concepts. In particular, Vrugt et al. (2008) see data assimilation as technique for inverse modeling, whereas van Leeuwen, Cheng, and Reich (2015) regard inverse modeling as a branch of data assimilation. We interpret data assimilation as a unifying approach to estimate system state, model parameters, and model structure, following Liu and Gupta (2007). In our application, we employ the EnKF as algorithm for data assimilation.

Data assimilation represents knowledge using probability density functions (pdfs) $p(X = x)$, where x is the realization of a random variable X. Interpreting model or system states $\boldsymbol{u}$ and observations of these states $\boldsymbol{d}$ as such realizations of random variables, we use Bayes' theorem to update an estimated system state based on observations,

$$p(\boldsymbol{u}|\boldsymbol{d}) = \frac{p(\boldsymbol{d}|\boldsymbol{u})p(\boldsymbol{u})}{p(\boldsymbol{d})}. \tag{4.1}$$

Here, $p(\boldsymbol{u})$ is the prior information of the state before the observations are taken and $p(\boldsymbol{d}|\boldsymbol{u})$ is the conditional likelihood of the observations as function of the state itself. Van Leeuwen, Cheng, and Reich (2015) interpret $p(\boldsymbol{d})$ as probability of a specific measurement right before the observation. They point out that this is a normalization factor which can be computed

from the likelihood and the prior,

$$p(\boldsymbol{d}) = \int p(\boldsymbol{d}|\boldsymbol{u})p(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}.$$

The resulting posterior $p(\boldsymbol{u}|\boldsymbol{d})$ is interpreted as optimal estimate of the true system state resulting from prior information and observations.

Algorithms for data assimilation are classified as "filters" and "smoothers." Filters apply a momentary update based on observations at a single time. Smoothers perform an update based on an extended interval of observations. This has the advantage of incorporating future observations into a state update and the disadvantage that fewer such updates can be performed. Apart from the fact that filters can be applied in real-time, their computational effort also scales linearly in the dimension of the state space and the number of observations. In soil hydrological applications, EnKFs have been shown to outperform ensemble smoothers because more frequent state updates ease the data assimilation task if the model is strongly non-linear (Crestani et al. 2013).

4.1 Ensemble Kalman Filter

The original Kalman filter assumes the random variables of states and observations to be normally distributed, which means that any state is characterized by a vector of mean values $\boldsymbol{u}$ and a covariance matrix

$$\boldsymbol{P} = \overline{[\boldsymbol{u} - \boldsymbol{u}_{\text{true}}]\,[\boldsymbol{u} - \boldsymbol{u}_{\text{true}}]^T}, \tag{4.2}$$

where the overline denotes the expectation value and $\boldsymbol{u}_{\text{true}}$ is the true state. The pdf of a state $\boldsymbol{u}$ and the corresponding covariance matrix $\boldsymbol{P}$ is given by a normal distribution $\mathcal{N}(\boldsymbol{u}, \boldsymbol{P})$,

$$p(\boldsymbol{x}) \propto \exp\left(-\frac{1}{2}\,[\boldsymbol{x} - \boldsymbol{u}]^T\,\boldsymbol{P}^{-1}\,[\boldsymbol{x} - \boldsymbol{u}]\right), \tag{4.3}$$

where $\boldsymbol{x} \in \mathbb{R}^{N_s}$ is the value of a vector of independent random variables $\boldsymbol{X}$, and N_s is the dimension of the state space. To receive the state pdf at a later time, a forecasting model has to be applied to the previous mean and covariance matrix, which is particularly simple if the model is linear and its application can be formulated as matrix (multiplication).

The EnKF is a Monte-Carlo extension of the Kalman filter for non-linear models (Evensen 1994, 2003). It represents the state pdf with an ensemble of states, and approximates the

first and second moments of this distribution with

$$\boldsymbol{u} \approx \overline{\boldsymbol{u}^i}, \tag{4.4}$$

$$\boldsymbol{P} \approx \overline{\left[\boldsymbol{u}^i - \overline{\boldsymbol{u}^i}\right]\left[\boldsymbol{u}^i - \overline{\boldsymbol{u}^i}\right]^T}, \tag{4.5}$$

where $i = 1, \ldots, N_{\text{ens}}$ denotes the ensemble member and N_{ens} is the ensemble size. From the perspective of the filter, the model is a "black-box" operator $M(\cdot)$ which takes a state $\boldsymbol{u}^i_{t-1}$ at time $t-1$ as argument and computes the state $\boldsymbol{u}^i_t$ at time t while adding a model error or "process noise" $\boldsymbol{\eta}^i$,

$$\boldsymbol{u}^i_t = M(\boldsymbol{u}^i_{t-1}) + \boldsymbol{\eta}^i. \tag{4.6}$$

The process noise is assumed to be normally distributed with mean 0 and error covariance $\boldsymbol{Q}$, $\boldsymbol{\eta} \sim \mathcal{N}(0, \boldsymbol{Q})$, and is intended to account for errors that are not represented in the model. Note that an initial ensemble of states $\boldsymbol{u}^i$ that follows a normal distribution will lose this property if the model operator is non-linear in $\boldsymbol{u}$.

The EnKF uses the model to propagate the ensemble states to the time t of the next observation $\boldsymbol{d}_t$. This observation is then compared with artificial observations of the states $\boldsymbol{u}^i_{t,f}$, where the subscript f indicates the state resulting from a model forecast. Like the original Kalman filter, the EnKF assumes an observation on a state to be a linear operation $\boldsymbol{H}_t$, resulting in the noisy observation

$$\boldsymbol{y}^i_t = \boldsymbol{H}_t \boldsymbol{u}^i_{t,f} + \boldsymbol{\epsilon}^i_t, \tag{4.7}$$

with the measurement error $\boldsymbol{\epsilon}^i_t$ being drawn from another unbiased normal distribution with covariance $\boldsymbol{R}_t$, $\boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \boldsymbol{R}_t)$. The perturbation by this error is required for an adequate comparison because the observations $\boldsymbol{d}_t$ are intrinsically perturbed by a measurement error (Burgers, Jan van Leeuwen, and Evensen 1998; van Leeuwen, Cheng, and Reich 2015). Each forecast state of the ensemble is then updated based on the difference between the observation $\boldsymbol{d}_t$ and the measurement of the state $\boldsymbol{y}^i_t$, yielding an analysis state

$$\boldsymbol{u}^i_{a,t} = \boldsymbol{u}^i_{t,f} + \boldsymbol{K}_t \left[\boldsymbol{d}_t - \boldsymbol{y}^i_t\right]. \tag{4.8}$$

The *Kalman gain* $\boldsymbol{K}$ provides a covariance-weighted map from the observation vector space to the state vector space,

$$\boldsymbol{K}_t \coloneqq \boldsymbol{P}_{t,f} \boldsymbol{H}_t^T \left[\boldsymbol{H}_t \boldsymbol{P}_{t,f} \boldsymbol{H}_t^T + \boldsymbol{R}_t\right]^{-1}, \tag{4.9}$$

where $\boldsymbol{P}_{t,f}$ denotes the covariance matrix estimated from the forecast states $\boldsymbol{u}^i_{t,f}$ according to eq. (4.5). Assuming an infinite ensemble size and a normal distribution of the forecast states, the EnKF is equivalent to the Kalman filter and its update minimizes the trace of the

error covariance $\boldsymbol{P}_{a,t}$ of the analysis ensemble, which is given by

$$\boldsymbol{P}_{a,t} = [\boldsymbol{I} - \boldsymbol{K}_t \boldsymbol{H}_t]\, \boldsymbol{P}_{t,f}. \tag{4.10}$$

The pdf of the resulting analysis ensemble is equivalent to the posterior pdf of Bayes' theorem (Bauser 2018),

$$p(\boldsymbol{u}_t|\boldsymbol{d}_t) \propto \exp\left(-\frac{1}{2}\left[\boldsymbol{u}_t - \overline{\boldsymbol{u}_{a,t}}\right]^T \boldsymbol{P}_{a,t}^{-1}\left[\boldsymbol{u}_t - \overline{\boldsymbol{u}_{a,t}}\right]\right). \tag{4.11}$$

The model is then applied on the ensemble of analysis states $u_{t,a}^i$ to yield the forecast state at the next observation time $t + 1$, thus concluding the iteration.

For model parameter estimation, the estimated parameters can be wrapped into a vector $\boldsymbol{\phi}$ and appended to the model state $\boldsymbol{\chi}$, yielding an *augmented state*

$$\boldsymbol{u} = [\boldsymbol{\chi}, \boldsymbol{\phi}]^T. \tag{4.12}$$

The model M_u for propagating an augmented state according to eq. (4.6) then has two separate components M_χ and M_ϕ for propagating the respective part of the augmented state. As model parameters are usually considered constant in time, the model M_ϕ for propagating them is a trivial identity map with $M_\phi(\boldsymbol{\phi}) = \boldsymbol{\phi}$, and the application of the combined model M_u yields

$$\boldsymbol{u}_{t+1,f} = M_u(\boldsymbol{u}_{t,a}) = \left[M_\chi(\boldsymbol{\chi}_{t,a}), M_\phi(\boldsymbol{\phi}_{t,a})\right]^T = \left[\boldsymbol{\chi}_{t+1,f}, \boldsymbol{\phi}_{t,a}\right]^T. \tag{4.13}$$

The EnKF assumes that the ensemble of states, the model error, and the measurement error are normally distributed and it updates the state ensemble based on this assumption. If a non-linear model is then applied on the analysis ensemble, the resulting forecast states will necessarily have a different distribution. Although the EnKF performance is directly related to the distribution of the ensemble, and especially if it can be approximated with a normal distribution, the filter performs successfully in many scenarios where this is not the case (Evensen 2003). As long as the state ensemble is not skewed too much, and there are many observations, the Gaussian model and measurement errors will continuously reduce higher moments in the state distribution. This probabilistic approach has been shown to outperform filter variants with analytic updates, like the ensemble square root filter, given that the ensemble size is large enough (Lawson and Hansen 2004).

Another method closely related to the EnKF is the particle filter (Leeuwen 2009). It considers higher moments of possibly multimodal state distributions and is especially suitable for low dimensional, chaotic systems. While it can be applied for data assimilation in soil hydrology, the particle filter requires sophisticated localization and resampling techniques to reach an accuracy and efficiency comparable to the EnKF (Berg 2018). Generally though, the particle filter requires a larger ensemble. The probabilistic update of the EnKF also avoids

the computation of adjoint or inverse matrices, and especially does not require the Jacobian of the model operator. This is the main advantage of the EnKF against variational methods. The relatively low computational cost of the EnKF makes it suitable for applications in soil hydrology, where the model operator contributes by far the most computational cost to the entire data assimilation task.

4.2 EnKF Extensions for Soil Hydrology

In practice, the limited ensemble size, non-linear models, and misrepresented or biased model errors significantly hamper the performance of the EnKF. These issues usually culminate in an underestimation of the state error covariance $\boldsymbol{P}$. In the extreme case $\boldsymbol{P} \to \mathbf{0}$, the Kalman gain becomes minimal and the analysis state is equivalent to the forecast state, ignoring any update by the observations according to eq. (4.8). As the observations are the crucial information supplied, many extensions of the EnKF focus on increasing the state error covariance. However, the small ensemble size also leads to spurious correlations, which need to be counteracted by a spatial localization.

Damping Through nonlinearities in the parameterization, the parameters defining the material properties have a profound influence on the dynamics of soil hydraulic systems. Since observations of these parameters are not possible, they are considered more "valuable" information than the soil hydraulic state, which is updated for every observation. To reduce the effect of nonlinear relations between soil hydraulic state and parameters in the linear state update, and also to reduce effects of spurious correlations, the state update can be damped with a constant factor $\gamma \in (0, 1]$ multiplied to the correction vector in eq. (4.8) (Hendricks Franssen and Kinzelbach 2008). This factor is usually extended to a vector $\boldsymbol{\gamma}$ with element-wise multiplication to enforce a stronger damping on parameters than on the soil hydraulic state (Wu and Margulis 2011).

Multiplicative Inflation A simple way of increasing the forecast covariance $\boldsymbol{P}_f$ is an artificial spread of the ensemble around the ensemble mean using an inflation factor $\lambda \geq 1$ to yield inflated ensemble states

$$\boldsymbol{u}^i_{f,\text{infl}} = \sqrt{\lambda} \left[\boldsymbol{u}^i_f - \overline{\boldsymbol{u}^i_f} \right] + \overline{\boldsymbol{u}^i_f}. \tag{4.14}$$

Like the damping factor, the inflation factor can be extended to a vector. Anderson (2009) optimized the inflation with a spatially and temporally adaptive formulation. In this application, we use the method by Bauser et al. (2018), which is specialized for the use in

soil hydrology and employs a Kalman filter to estimate the inflation for every observation independently.

Localization Spurious correlations are noise in the covariance matrix resulting from a reduced ensemble size and a therefore non-optimal estimation according to eq. (4.5). Some of these spurious correlations can be cut off based on the observation that correlations in states at two different locations decrease with increasing distance (Hamill, Whitaker, and Snyder 2001). The covariance matrix can therefore be multiplied with a correlation function ρ based on the physical distance between state dimensions (Houtekamer and Mitchell 2001), altering the Kalman gain definition of eq. (4.8) to

$$\boldsymbol{K}_t = \left[\rho \circ \boldsymbol{P}_{t,f}\right] \boldsymbol{H}_t^T \left[\boldsymbol{H}_t \left[\rho \circ \boldsymbol{P}_{t,f}\right] \boldsymbol{H}_t^T + \boldsymbol{R}_t\right]^{-1}, \tag{4.15}$$

where $\circ$ denotes the product of the result of ρ with the respective entry in the covariance matrix.

4.3 Data Assimilation with DORiE

The Richards solver of DORiE was coupled with the research group internal data assimilation software package Knowledge Fusion (KnoFu; Berg 2018) to estimate soil hydraulic states and parameters using a DG solver (Riedel 2017). This coupling has been upgraded throughout the development of DORiE and now includes the FV solver and the option to estimate scaling factors. KnoFu handles the application of the filter algorithm. For the forecast between measurements, the filter states are passed to the model, which then propagates the state to the next measurement time and returns the new state. To that end, KnoFu initializes an instance of the DORiE solver that works like an independent instance of DORiE, with the difference that it can be called through an application programming interface (API). This API also supplies model specific information which is required by an EnKF specialized for soil hydrology as outlined by Bauser et al. (2018).

Model State The unknown variable within DORiE is the matric head h_m and selecting it as state variable in the filter would be a natural choice. However, the EnKF requires a linear observation operator $\boldsymbol{H}$ and soil hydraulic state observations often supply water content data, e.g. from time domain reflectometry (TDR) probes. If the observations are water content values, DORiE therefore transforms its internal state to water content values via the soil water characteristic, eq. (1.13), before passing it to the filter as forecast state. Likewise, the initial state received from the filter is transformed back into matric head values. This transformation is unique in an unsaturated regime where $h_m \leq 0$, if local equilibrium without

hysteresis is assumed. Using water content values is generally considered advantageous because the skewness of the state distribution caused by extreme matric head values is reduced under the transformation (Erdal, Rahman, and Neuweiler 2015).

Localization DORiE returns physical positions associated with each dimension of the state vector to the filter. In case of the FV discretization, the position is the barycenter of the grid cell the respective dimension in the state vector is associated with. In contrast, the solution at every point of the domain in a DG function space is given by a linear combination of basis functions and DOFs on the local element. However, a local function space with Langrange polynomials as basis features nodes—i.e., positions on the grid—where the value of the solution is given by exactly one DOF. Since the DUNE-Localfunctions module does not supply an interface for such information, the positions for common geometries and polynomial orders were hard-coded and the associated state dimension is selected by searching for the local basis function whose value is sufficiently close to 1.

Discretization Scheme Selection Employing an EnKF with a DG Richards solver has proven difficult because the solver is often unable to propagate the filter analysis states, especially when parameters are estimated (Riedel 2017). This can be attributed to the instability of the DG method and its susceptibility for over- and undershoots, as discussed in section 3.2.2. The forecast state is a consistent result of the numerical method, but the analysis state is the result of a probabilistic update in the filter algorithm and not necessarily physically consistent. While localization helps in reducing these inconsistencies, it does not apply to the parameters. Their update, in turn, can imply large jumps and gradients in the matric head after transforming the water content state received by the filter. The issue can be reduced by increasing the penalty factor m, which penalizes solution jumps $[[u_h]]$ across interfaces, cf. eqs. (2.14d) and (2.17). Additionally, the grid resolution can be increased, but this also increases the computational cost of the forward model. Overall, a stable, less accurate numerical method is preferable to an unstable, more accurate one because the variance of the state ensemble is orders of magnitude larger than the discrepancy between, e.g., FV and DG discretization methods, at least for high grid resolutions.

5 Effects of Unrepresented Soil Heterogeneity

This chapter is based on, and extends, Bauser et al. (2020): "Challenges with effective representations of heterogeneity in soil hydrology based on local water content measurements."

5.1 Research Question

With the highly non-linear effects of material properties on soil hydraulic fluxes presented in chapter 1, these material properties are the primary source of uncertainty when estimating and forecasting soil hydraulic states. Soils have been identified as complex systems featuring heterogeneity at every spatial scale, and efforts to investigate self-similarity of different scales are ongoing and not yet conclusive (Pachepsky and Hill 2017). Although the scale of interest for many problems in hydrology is larger, the spatial scale of soil profiles ranging from centimeters to a few meters has been identified as the critical scale for soil hydrology, governing larger-scale processes (Vogel 2019). Even on this scale, the inherent heterogeneity makes it difficult to deduce material properties for extended soil layers. This is aggravated by the fact that soil hydraulic states are difficult to observe at high spatial resolutions or high accuracy.

Soil hydraulic parameters can be estimated from in situ water content or matric potential measurements using inverse modeling and data assimilation (Vrugt et al. 2008; Liu et al. 2012). If the soil is known to have distinct layers, independent parameter sets can be estimated for each layer (Botto, Belluco, and Camporese 2018). Incorporating heterogeneity within such layers is more difficult. To account for local deviations of the water content, Bauser et al. (2016) estimated an additional scaling factor at each measurement location and combined these linearly to a scaling field. This field yielded an effective description of the true heterogeneity and replicated the observed dynamics at the measurement locations. If the measurement density is high enough, data assimilation can also be used to estimate an entire three-dimensional heterogeneity field (Chaudhuri, Hendricks Franssen, and Sekhar 2018). However, any such estimation is limited by the spatial resolution of the measurements and does not account for heterogeneity at a lower scale. A different approach is therefore to not estimate heterogeneity itself but to acknowledge unresolved heterogeneity as structural model error and estimate its effect on the model states through a data-driven Gaussian process regression (Zhang et al. 2019).

The effect of unresolved heterogeneity is difficult to quantify, especially prior to experiments. In the presented study, we aimed to investigate the isolated effects of unrepresented heterogeneity on the accuracy of a forecast performed with estimated soil properties. Our synthetic experiments consisted of computing a synthetic truth on a two-dimensional domain with small-scale heterogeneity and estimating this domain by using a one-dimensional model with reduced representation of heterogeneity. In both cases, however, heterogeneity was achieved by a Miller scaling field, and the medium contained only a single reference material. We followed the approach of Bauser et al. (2016) to only estimate scaling factors at measurement locations, but added a scaling factor to be estimated at the surface. These estimated factors were combined to a field, which, by design of the experiment, could not resolve the heterogeneity of the synthetic truth. The estimated model therefore accounted neither for lateral fluxes induced by heterogeneity, nor the full heterogeneity itself.

5.2 Synthetic Experiments

The presented study conducted 80 synthetic experiments including unrepresented, small-scale soil heterogeneity as structural model error. We chose to simulate a two-dimensional, $2\,\mathrm{m} \times 2\,\mathrm{m}$ domain with four local water content measurements at its center in depths of 9.5 cm, 19.5 cm, 39.5 cm and 79.5 cm. While focusing on varying heterogeneity, we used two soil types and two boundary condition sequences to indicate that the results are generalizable to a certain extent. In this thesis, I will limit my analyses to a single medium and a single boundary condition sequence. This reduces the investigated scenarios to the ones listed in table 5.1, comprising a total of 50 experiments.

5.2.1 Model Setup

The reference material for all scenarios considered in this thesis is *Sandy Loam*, cf. table 1.1. The small-scale heterogeneity is created by employing Miller scaling, see section 1.5. For every scenario, we generated five realizations of a two-dimensional, scalar Gaussian random field with mean value 0. A control case for each realization with effectively one-dimensional scaling field was created by taking a vertical slice at the center of the field. We used this control case to investigate unrepresented heterogeneity without the influence of two-dimensional flow. We created the Gaussian random fields with a resolution of 1 cm using the random field generator included in DORiE (Klein 2016). The autocovariance function used was Gaussian and the anisotropy was defined by the vertical and horizontal correlation lengths ℓ_v and ℓ_h. The Miller scaling factors at position $\boldsymbol{x}$ were then calculated from the

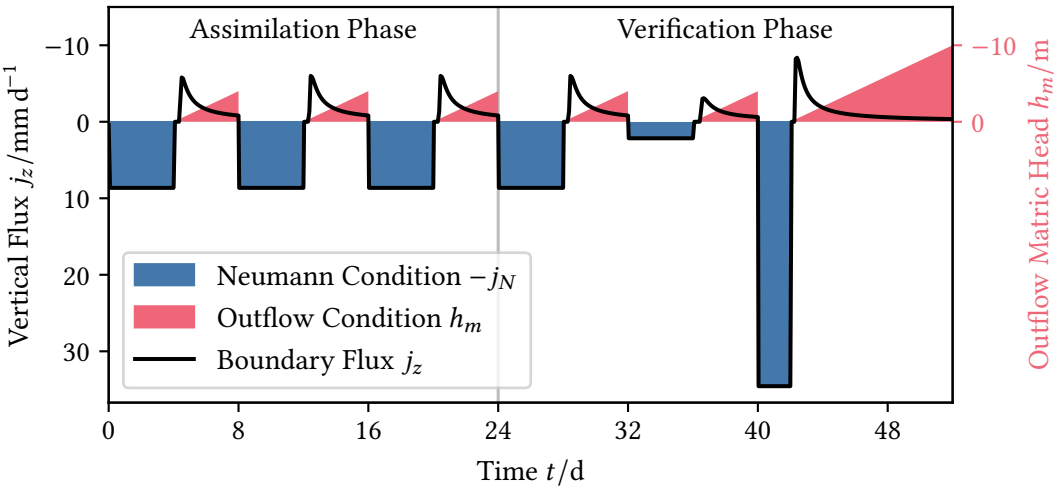

Figure 5.1: Boundary condition sequence applied in the data assimilation cases discussed in chapter 5. Shaded areas denote the (negative) Neumann flux $-j_N$ (blue, left y-axis scale) and outflow matric head h_m (red, right scale) applied at the upper boundary. The black line (left scale) indicates the resulting vertical flux j_z at the upper boundary of a homogeneous *Sandy Loam* domain with the model setup described in section 5.2.1. Note that a positive vertical flux points downwards, but a positive Neumann flux at the upper boundary points upwards.

value of the random field $r(\boldsymbol{x})$ and the prescribed field variance σ_ξ according to

$$\xi(\boldsymbol{x}) = \exp\left(r(\boldsymbol{x}) - \sigma_\xi^2\right), \tag{5.1}$$

which keeps the expectation value of the resulting hydraulic conductivity equal to that of the unscaled reference material (Roth 1995).

The initial condition for all experiments was created through a 30 d spin-up simulation. The initial condition for this simulation was hydraulic equilibrium, $\partial_z h_m = 1$, with a water table at the lower end of the domain in a depth of 2 m. This water table also was the fixed lower boundary condition for the entire simulation. For the first 10 d, we simulated a constant infiltration of $j_N = -10^{-7}\,\mathrm{m\,s^{-1}}$, followed by a no-flow boundary condition for the rest of the simulation. The final state of the simulation was used as initial condition for the synthetic experiment.

The total simulated time of each experiment was 52 d. This time was split in two phases, (i) an *assimilation phase* of 24 d, where an EnKF estimates the soil hydraulic state, material properties, and small-scale heterogeneity, and (ii) a *verification phase* of 28 d, where the

prediction capabilities of the estimated material properties were tested with an ensemble forecast.

Each phase consisted of three alternating irrigation and evaporation events (Figure 5.1). During the assimilation phase, each rain event lasted for 4 d with a flux of $j_N = -10^{-7}\,\mathrm{m\,s^{-1}} \approx -8.64\,\mathrm{mm\,d^{-1}}$. The same flux was applied during the first irrigation event of the verification phase. The next event had a reduced influx of $2.5 \times 10^{-8}\,\mathrm{m\,s^{-1}}$ ($2.68\,\mathrm{mm\,d^{-1}}$), and the last rain event an increased influx of $4 \times 10^{-7}\,\mathrm{m\,s^{-1}}$ ($34.56\,\mathrm{mm\,d^{-1}}$), but the latter only lasted for 2 d. All evaporation events were modeled with an outflow boundary condition which avoids a flux into the domain. Its value was chosen to linearly increase by $1\,\mathrm{m\,d^{-1}}$ starting with a potential of $h_m = 0\,\mathrm{m}$ at the end of the preceding irrigation event. The first five evaporation events lasted 4 d each, and the final event was extended to 10 d. The lower boundary condition at a depth of 2 m was a constant water table, and no-flow boundary conditions were applied at the lateral domain boundaries. The rain events were chosen large and long enough to cause a response in all measurement depths.

The model structural error was introduced by using different models for generating the synthetic measurements and for propagating the ensemble states during the data assimilation. The state ensemble was propagated by an effectively one-dimensional model with a vertical resolution of 1 cm and a single grid cell in the horizontal direction. Its Miller scaling field was reconstructed from five estimated scaling factors, as explained in the following subsection. The synthetic measurements were generated by using models with fully resolved scaling field. The two-dimensional model used 200×200 grid cells. For the control case, the one-dimensional model was used, and the applied scaling field was a slice of the two-dimensional scaling field at the center. In the following, we will refer to the former as *2D* case (two-dimensional synthetic truth with one-dimensional estimation model) and to the latter as *1D* case (one-dimensional synthetic truth with one-dimensional estimation model). This results in two experiments for each random field realization. We decided to use the FV discretization because it is computationally cheaper and more stable than the DG scheme, cf. section 4.3.

Reduced Miller Field Representation

From the set of estimated Miller scaling factors $\boldsymbol{\xi}$ at certain locations in the domain, a scaling field $\xi(z)$ at depth z was reconstructed following

$$\xi(z) = 1 + \boldsymbol{\varphi}(z)^T \boldsymbol{A}(z)^{-1} \left[\boldsymbol{\xi} - \mathbf{1}\right], \tag{5.2}$$

where $\boldsymbol{\varphi}$ is a vector containing one kernel function $\varphi_i(z)$ for each scaling factor ξ_i, and $\boldsymbol{A}(z)$ is the local correlation matrix of the scaling factors. The kernel functions are chosen to be unnormalized Gaussian functions whose means are the locations of the estimated

scaling factors, respectively, and whose standard deviation is 10 cm. The correlation matrix $\boldsymbol{A}$ consists of the products of these kernel functions,

$$A_{ij}(z) = A_{ji}(z) = \varphi_i(z)\varphi_j(z), \quad A_{ii} = 1. \tag{5.3}$$

This heuristic choice makes the estimated scaling factors the extrema or saddle points of the resulting field. With increasing distance to the estimation locations, the scaling relaxes to the reference material properties, $\xi(z) \to 1$. See section 5.2.2 for the estimation locations.

Apart from the choice of Gaussian functions, the only prior information required for this method is the standard deviation used in the kernels. The value is ideally set to an estimated correlation length of the true heterogeneity. However, the method is expected to be insensitive against the exact value of the kernel standard deviation used, as long as its magnitude is reasonably approximated.

5.2.2 Data Assimilation

The synthetic measurements of the water content at the given locations were generated every 2 h and perturbed with an unbiased normally-distributed measurement uncertainty of 0.007, which is a typical value for TDR probes measuring water content in situ (Jaumann and Roth 2017). During the assimilation phase, these observations were used as input for the EnKF to estimate the one-dimensional soil hydraulic state, the Mualem–van Genuchten parameters α, $\lg K_0$, and n, the Miller scaling factors $\lg \xi$ at the measurement locations, and an additional Miller scaling factor at the top grid cell. The scaling factors were therefore estimated in depths of 0.5 cm, 9.5 cm, 19.5 cm, 39.5 cm and 79.5 cm. As indicated, the decimal logarithm of the saturated hydraulic conductivity K_0 and the scaling factors ξ were used in the data assimilation. The parameters θ_r, θ_s, and τ were set to their true values. In the verification phase, the EnKF was turned off and the model only produced an ensemble forecast.

The initial guess for the estimated water content was a linear interpolation of the synthetic measurements of the initial condition. The values towards the upper boundary were extrapolated from the topmost measurement, and the values towards the lower boundary were linearly extrapolated towards the saturated water content θ_s, which corresponds to the lower boundary condition. The initial uncertainty of the ensemble was created with an unbiased normal distribution with standard deviation 0.01 and a spatial covariance given by an approximated normal distribution with standard deviation 20 cm. The initial guess for the Mualem–van Genuchten parameters and the Miller scaling factors were normal distributions, with their mean set to the respective true value to reduce the impact of the EnKF performance. The standard deviations were $\sigma_\alpha = 2.0\,\mathrm{m}^{-1}$, $\sigma_n = 0.25$, $\sigma_{\lg K_0} = 1.0$, and $\sigma_{\lg \xi} = 0.2$.

Table 5.1: Overview and results of experiments conducted by Bauser et al. (2020) for selected scenarios. The center column group lists the properties of the generated Miller scaling fields. The right column group contains the averaged results of the experiments. The overline denotes the mean calculated from the five cases run for each combination.

Scenario	Miller Scaling Field			$\overline{\text{AHFR}}$	$\overline{\text{RMSE}_\theta}$		$\overline{\lvert\Delta J\rvert}$		$\overline{\text{NSE}_j}$	
	σ_ξ^2	ℓ_h/cm	ℓ_v/cm	2D	1D	2D	1D	2D	1D	2D
Default	0.1	10	10	0.14	0.0022	0.0040	0.02	0.10	0.97	0.84
MedVar	0.2	10	10	0.18	0.0029	0.0049	0.03	0.13	0.92	0.73
HighVar	0.5	10	10	0.24	0.0051	0.0072	0.05	0.15	0.78	0.36
Horizontal	0.1	30	10	0.09	0.0021	0.0022	0.02	0.05	0.98	0.97
Vertical	0.1	10	30	0.09	0.0019	0.0041	0.01	0.06	0.97	0.92

5.2.3 Evaluation

We characterized and evaluated the predictions from the verification phase of each experiment. As indicator for the relevance of lateral, two-dimensional fluxes we calculated the average horizontal flux ratio (AHFR) from the spatial average of the absolute horizontal flux in the top 100 cm of the domain, $\langle|j_{x,t}|\rangle_z$, and the spatial average of the vertical flux in the same region, $\langle|j_{z,t}|\rangle_z$, as

$$\text{AHFR} = \frac{1}{N_t}\sum_{t}^{N_t}\frac{\langle|j_{x,t}|\rangle_z}{\langle|j_{z,t}|\rangle_z}, \tag{5.4}$$

where t indexes the time step with respect to the measurements, and $N_t = 28\,\text{d}/2\,\text{h} = 336$ is the total number of observations in the verification phase. This value could only be obtained for the two-dimensional synthetic truth, for all other models it would be 0. We therefore use the value from the two-dimensional synthetic truth to characterize both the 1D and the 2D case from the same scaling field realization.

The water content root mean squared error (RMSE) is calculated by

$$\text{RMSE}_\theta = \sqrt{\frac{1}{N_t N_m}\sum_{t}^{N_t}\sum_{m}^{N_m}\left[\theta^*_{t,m} - \overline{\theta_{t,m}}\right]^2}, \tag{5.5}$$

where the index m denotes the measurement location a total number of measurements per observation $N_m = 4$, and where $\theta^*_{t,m}$ and $\overline{\theta_{t,m}}$ are the true water content and the ensemble mean water content at measurement location m and observation index t, respectively.

The flux prediction is evaluated with two different measures. We calculate the relative

cumulative flux error at a depth of 100 cm as

$$\Delta J = \frac{\sum_t^{N_t} \left[\overline{j_{z,t}} - \left\langle j_{z,t}^* \right\rangle_x \right]}{\sum_t^{N_t} \left\langle j_{z,t}^* \right\rangle_x}, \tag{5.6}$$

where $\overline{j_{z,t}}$ and $\left\langle j_{z,t}^* \right\rangle_x$ are the ensemble mean and the horizontally averaged true vertical flux for observation t, respectively. The Nash-Sutcliffe efficiency (NSE) of the flux is calculated by

$$\mathrm{NSE}_j = 1 - \frac{\sum_t^{N_t} \left[\overline{j_{z,t}} - \left\langle j_{z,t}^* \right\rangle_x \right]^2}{\sum_t^{N_t} \left[\left\langle j_{z,t}^* \right\rangle_x - \sum_t^{N_t} \left\langle j_{z,t}^* \right\rangle_x / N_t \right]^2}. \tag{5.7}$$

This measure indicates the quality of the flux prediction compared to the temporally averaged (mean) flux. All fluxes are evaluated at a depth of 100 cm.

5.3 Results and Discussion

We find that the EnKF is able to estimate effective material properties which adequately reproduce the dynamics of the fully resolved scaling field for most cases (Table 5.1 and fig. 5.2). The error of the water content at the measurement locations, RMSE_θ, grows with the heterogeneity of the local flow field indicated by AHFR, and is usually higher for the respective 2D case, but generally lies below a value of 0.01 and therefore in the range of the measurement uncertainty of TDR probes. Fluxes can also be estimated well, although there are several severe outliers and no clear trend can be discerned. The relative cumulative flux error ΔJ of most cases lies within 10 %, but we also find several outliers with an error of up to 50 %, apparently independent of the scenario. The range of values for ΔJ slightly increases with increasing AHFR. In most cases, the cumulative flux is underestimated. The results for the NSE are similar. Most cases achieve an efficiency of $\mathrm{NSE}_j > 0.8$, indicating that the estimated model approximates the flux in a depth of 100 cm much better than the mean observed flux. The efficiency of 1D cases is usually slightly higher than that of the respective 2D cases. Overall, it decreases with increasing AHFR, and several outliers are close to, or even below, a value of $\mathrm{NSE}_j = 0$ that indicates the same predictive capabilities as the mean observed flux throughout the entire verification phase.

The *Horizontal* and *Vertical* scenarios were used to investigate different structures of the Miller similar media. Both scenarios feature the lowest mean AHFR. In case of *Horizontal*, the two-dimensional medium can locally be approximated well with a vertically layered, effectively one-dimensional medium. These experiments therefore yield the lowest observed errors throughout all considered scenarios. In case of *Vertical*, mostly vertical flow channels form through regions with high conductivity. The data assimilation produces reasonable

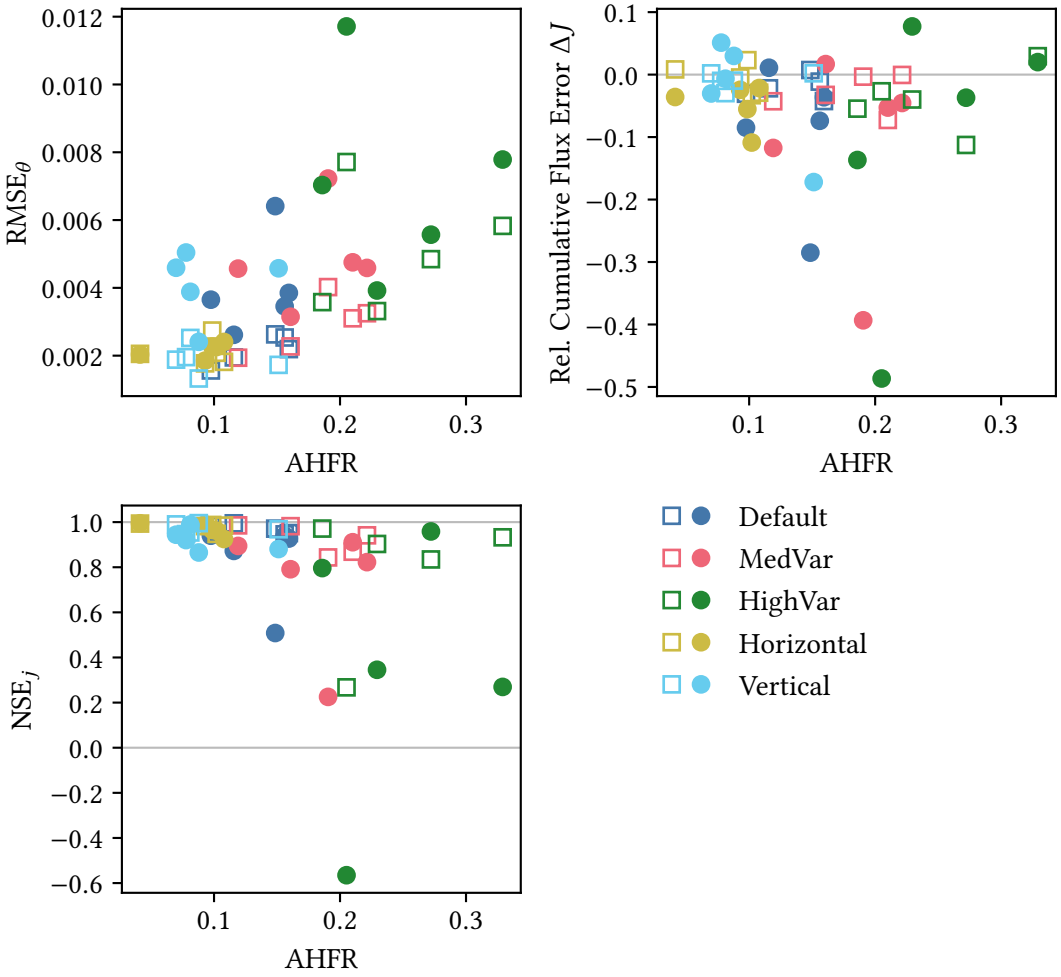

Figure 5.2: Summary measures for selected scenarios of experiments conducted by Bauser et al. (2020). The water content root mean squared error RMSE$_\theta$, relative cumulative flux error ΔJ, and flux Nash-Sutcliffe efficiency NSE$_j$ are plotted against the average horizontal flux ratio AHFR of the respective 2D case, which measures the heterogeneity of the flow field. Therefore, there are two values plotted for each case and AHFR: Rectangles denote the values for the 1D cases and circles denote the values for the 2D cases. Colors indicate the scenario, with scenario settings listed in table 5.1. Gray lines highlight the ideal values $\Delta J = 0$ and NSE$_j = 1$, and the critical value NSE$_j = 0$ below which the mean flux is a better description of the real flux than the estimated flux.

results if the soil profile does not match a single channel exactly, but if a representative amount of lateral fluxes between the channels can be observed.

In the following, I extended the analysis of the two cases the study by Bauser et al. (2020) focused on to emphasize the results.

5.3.1 Two-Dimensional Flow

The realization of the *HighVar* scenario with AHFR = 0.33 has the largest AHFR observed throughout all conducted experiments. Although the variance of the true Miller scaling field is high with $\sigma_\xi = 0.5$, and the average horizontal flux is one third of the average vertical flux, the error in the water content prediction is rather low with $\mathrm{RMSE}_\theta = 0.0058$ for the 1D case and $\mathrm{RMSE}_\theta = 0.0078$ for the 2D case. The relative cumulative flux error is also low with $\Delta J = 3\,\%$ for 1D and $\Delta J = 2\,\%$ for 2D, but the NSE indicates a much worse predictive capability of the flux with $\mathrm{NSE}_j = 0.93$ in 1D against a value of 0.27 in 2D.

In the 1D case, the water content predictions at all measurement locations follow the true dynamics well, apart from the lowest measurement, where the water content change caused by the last infiltration event is underestimated (Figure 5.3). In the 2D case, there are larger errors during the first infiltration event, and the arrival time of the final infiltration front is misjudged at all measurement locations, resulting in sharp spikes in the water content residual. Although the ensemble spread is slightly increased in the 2D case, both cases underestimate the uncertainty in the water content state, with the ensemble spread not covering the synthetic truth as structural model errors manifest through increased residuals.

The estimation results can be better understood when comparing the true and estimated Miller scaling fields. In the 1D case, the topmost part of the estimated scaling field is very similar to the true scaling field (Figure 5.4, top panels). This is due to the high density of measurements and estimated scaling factors in a depth from 0 cm to 20 cm. Between the third and the fourth measurement location (counting from the top), the true scaling field features two layers with high scaling factor and one layer with low scaling factor, which cannot be resolved by the reduced representation of the estimated field. The EnKF compensates by selecting an approximately mean scaling factor in depths below 40 cm. The resulting scaling field is able to estimate the arrival time of infiltration fronts well. As this mean value matches the true scaling factor at the third measurement location well, the water content prediction at this measurement works. However, the estimated value at the fourth measurement location is higher than the true value, indicating a more coarse-grained material than in the true field. Compared to a fine-grained material, we expect a lower water content for the same flux magnitude in a coarse-grained material. This is precisely

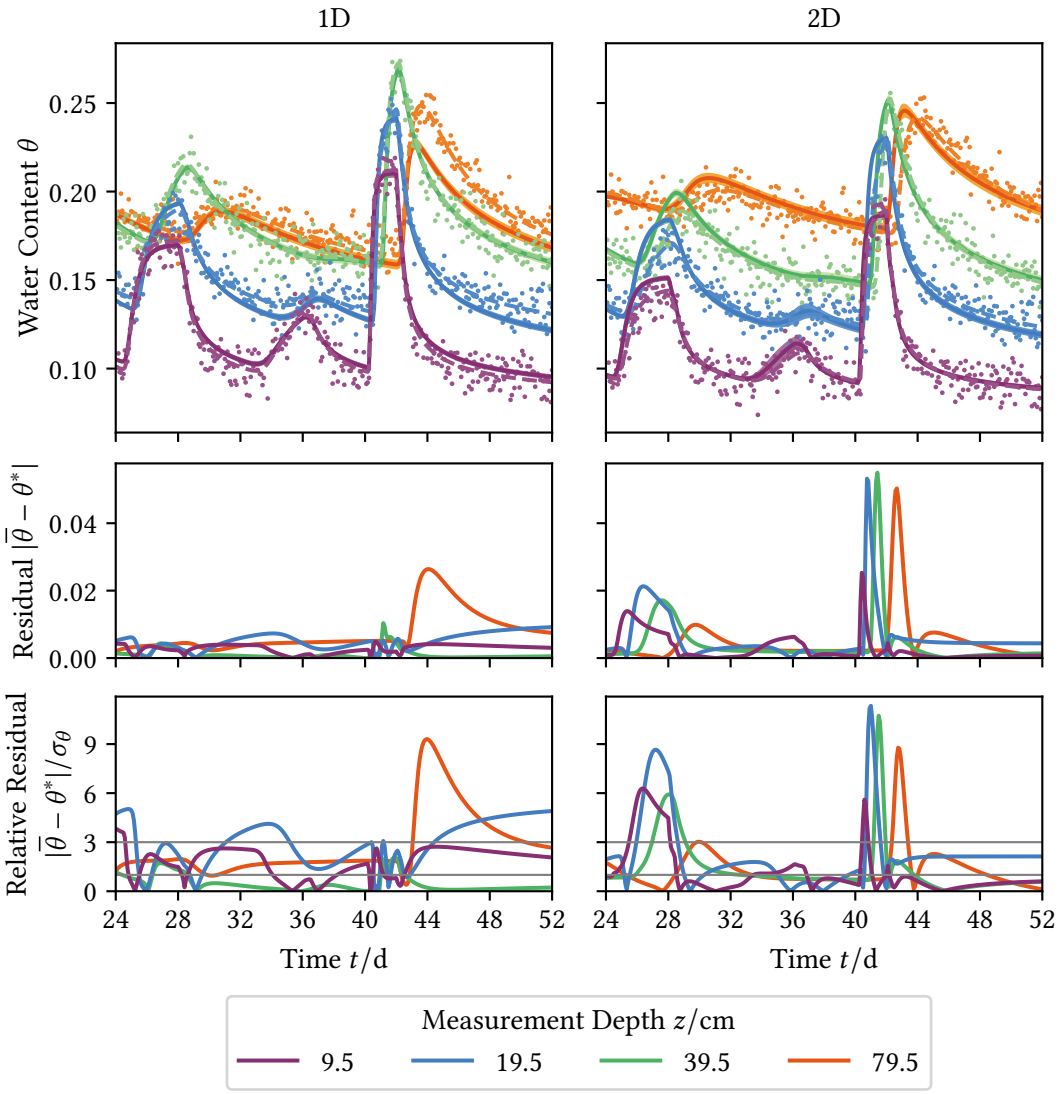

Figure 5.3: Water content predictions at the four measurement depths for the 1D and 2D cases of the *HighVar* scenario with the largest AHFR value. The top panels show the mean of the ensemble forecast (dark lines) with respective standard deviation (shaded areas), the synthetic truth (dashed lines), and the noisy measurements (markers). The center panels display the absolute residuals between the ensemble mean $\overline{\theta}$ and the synthetic truth θ^*, and the lower panels show the same values divided by the respective standard deviation σ_θ computed from the ensemble at this location. The gray lines indicate the σ_θ- and $3\sigma_\theta$-uncertainty ranges.

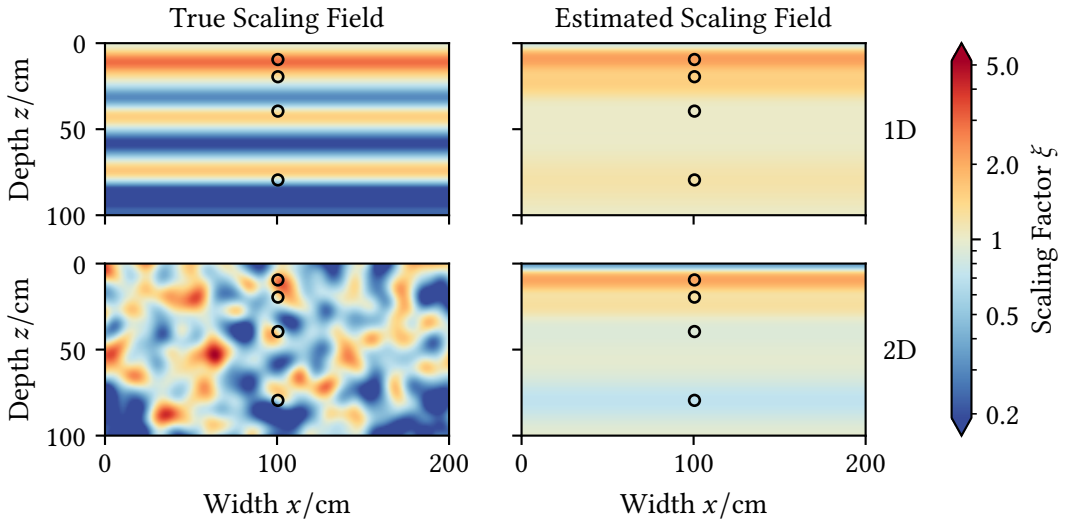

Figure 5.4: True and estimated Miller scaling fields for the 1D and 2D cases of the *HighVar* scenario with the largest AHFR value. Measurement positions are indicated by black circles and coincide with scaling factor estimation locations. One additional scaling factor is estimated at a depth of 0.5 cm.

what becomes visible in fig. 5.3, where the water content for the deepest measurement is underestimated, but the arrival time of the infiltration front is estimated well.

In the 2D case, the estimated scaling factors at the measurement locations mostly follow the true scaling factors at these locations (Figure 5.4, bottom panels). This improves the estimation of the local retention curve and thus the water content estimations when the system equilibrates after an infiltration has stopped. However, it worsens the estimation of the dynamics of the infiltration events themselves. The estimated scaling field does not account for the fine-grained regions in between the measurement locations, which cause a channelling and lateral deviations of strong infiltration fluxes. As a consequence, the propagation speed of the front is underestimated. Nonetheless, the evaporation from the topmost measurement location is correctly estimated by a reduced scaling factor near the surface. This does not represent the coarse-grained region directly above the measurement, but the more fine-grained regions next to it, which support a higher evaporation flux. The resulting effective representation is able to predict water contents and cumulative fluxes reasonably well, but, as indicated by the low NSE value, the flux estimation only reaches slightly better results than the mean flux.

The preceding arguments focus on the estimated scaling field for a description of the effective model dynamics. But the estimated saturated hydraulic conductivity K_0 and the pore size distribution n also have a profound effect on the dynamics. However, one can argue that the

scaling factors have a more immediate effect on the dynamics around the locations where they are estimated. Changes to the saturated hydraulic conductivity have less effect the lower the fluxes are—if the domain were in hydrostatic equilibrium, the saturated hydraulic conductivity could not be estimated. The pore size distribution n directly affects the relation between water content and matric head, but as a global parameter it also applies on the entire domain. While the scaling factors similarly affect the retention curve and the dynamics, they only have a localized effect. The latter allows them to account for unrepresented dynamics around their estimation location, but any deviation from the true scaling factors is immediately penalized because of their alteration of the local retention curve, at least as long as there is no global bias. It is therefore reasonable to assume that differences between true and estimated scaling fields are due to unrepresented heterogeneity with dynamics that cannot be explained by adjustments to the state or the global parameters K_0 and n.

5.3.2 Surface Heterogeneity and Evaporation

We further investigate the 2D case of the *HighVar* scenario with the highest RMSE and the worst NSE of all conducted experiments. The water content residuals are high during infiltration events, and the relative residual reaches values of more than six times the ensemble spread σ_θ (Figure 5.5, left panels). This is a prime indicator for structural model errors which are not represented by the ensemble uncertainty. We also note that the water content at the lowest measurement position only reacts significantly to the last, strong infiltration event. While the evaporation fluxes are overestimated, the infiltration fluxes are underestimated, and the predicted dynamics generally lag the observations (Figure 5.6, left panels).

The bad predictive performance is the result of a particular realization of the Miller scaling field. Right above the topmost measurement, the true scaling field reaches a very low scaling value of $\xi = 0.09$, inducing a very fine-grained medium. Additionally, the measurement is bordered by regions with large scaling factors in the horizontal direction, inducing coarse-grained media. The fine-grained medium at the top causes a strong evaporation flux, which is not representative of the mean evaporation flux, and the coarse-grained regions cause a funnel-like merger of infiltration fluxes towards the top measurements (Figure 5.7).

We created an additional experimental case by running the 2D case without estimating the Miller scaling factor in the surface grid cell. In this *modified 2D* case, the estimated scaling field was thus relaxed towards the surface. As the correlation lengths of the field are much lower than its extensions, this better resembles the average material properties of the synthetic truth at the surface. The modification increases the RMSE to a value of 0.014 due to fewer DOFs available to the EnKF for state optimization, but it also increases the NSE to 0.64, reduces the cumulative flux error to 19 %, and increases the overall ensemble uncertainty. Although the relative residuals still hint at structural model errors, they are

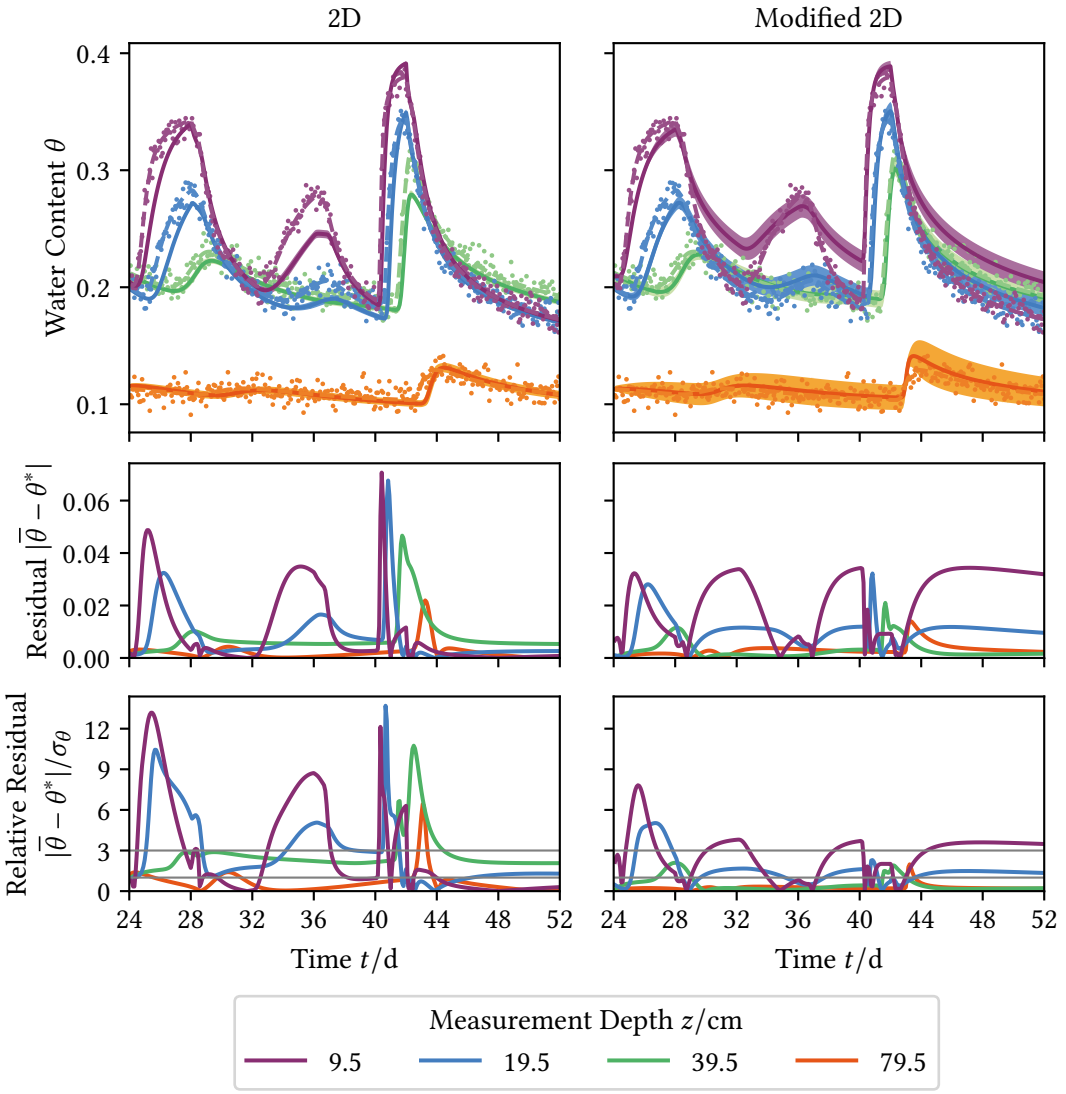

Figure 5.5: Water content predictions at the four measurement depths for the 2D case of the *HighVar* scenario with the worst prediction results (left) and for the same case with the modification of not estimating the topmost Miller scaling factor (right). The top panels show the mean of the ensemble forecast (dark lines) with respective standard deviation (shaded areas), the synthetic truth (dashed lines), and the noisy measurements (markers). The center panels display the absolute residuals between the ensemble mean $\overline{\theta}$ and the synthetic truth θ^*, and the lower panels show the same values divided by the respective standard deviation σ_θ computed from the ensemble at this location. The gray lines indicate the σ_θ- and $3\sigma_\theta$-uncertainty ranges.

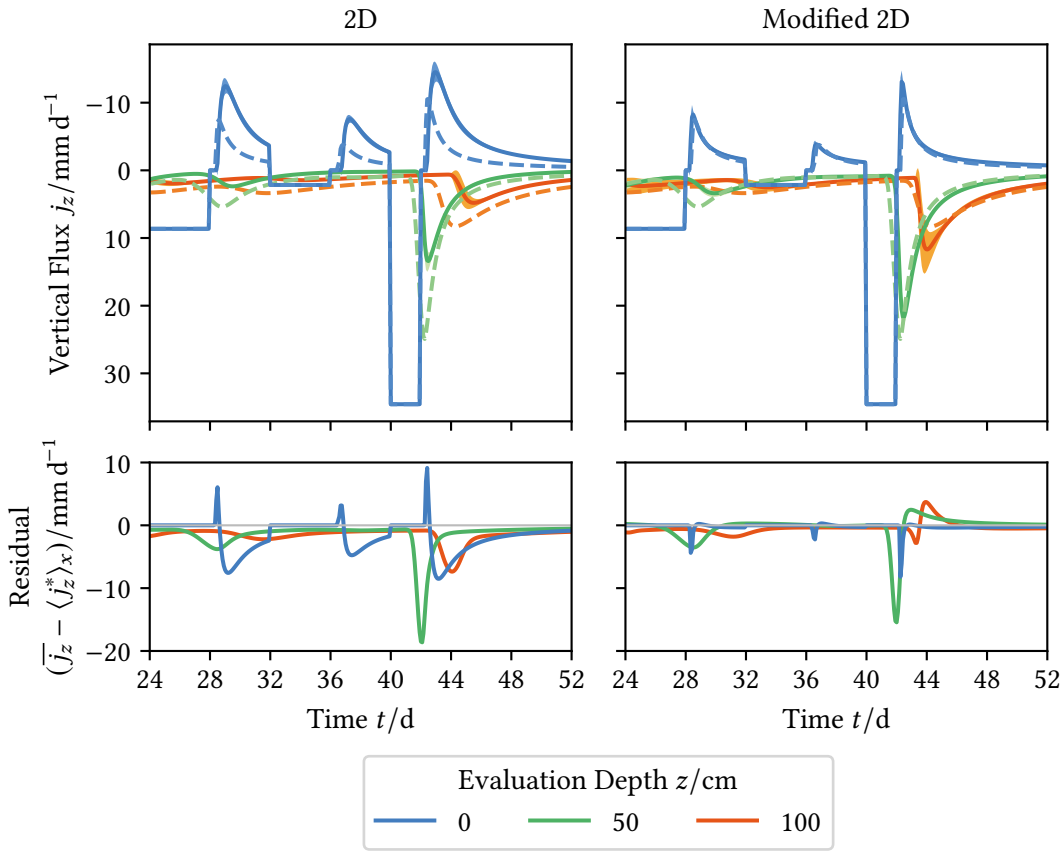

Figure 5.6: Flux predictions at three evaluation depths for the regular and modified 2D case of the *HighVar* scenario with the worst prediction results. The top panels show the mean of the ensemble forecast (dark lines) with respective standard deviation (shaded areas) and the synthetic truth $\langle j_z^* \rangle$ (dashed lines), which is the horizontal average of the vertical flux at the respective depth. Note that the vertical axis is inverted: a positive vertical flux component j_z indicates a downward flux. The lower panels display the residual between the ensemble mean $\overline{j_z}$ and the synthetic truth j_z^*.

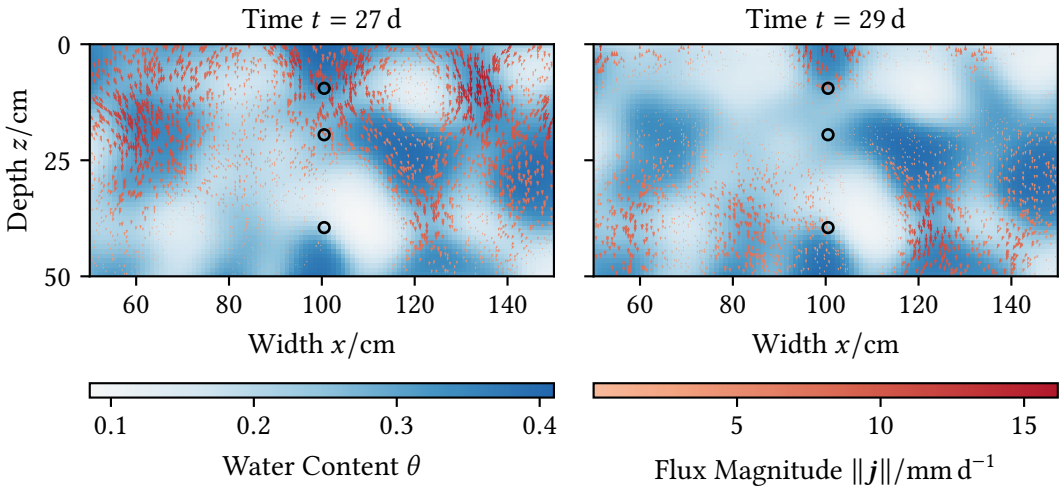

Figure 5.7: Water content and flow field of the synthetic truth at two different times t for the case of the *HighVar* scenario with the worst prediction results. Measurement positions are indicated by black circles. Arrows indicate flux direction and magnitude. The color bars apply to both panels.

significantly reduced (Figure 5.5, right panels). The flux prediction in all depths also clearly improves (Figure 5.6, right panels).

In the 2D case, the estimated water content better resembles the true water content above the profile (Figure 5.8). The state uncertainty grows towards the surface because there are no water content measurements available for the estimation of the scaling factor and the observed high evaporation flux can be explained by either a large water content or a low scaling factor. Contrary, through the scaling field relaxation, the estimated state of the modified 2D case better estimates the average water content at the surface. This can be considered coincidental: in the modified 2D case, the observed evaporation flux demands a stronger gradient in the water content because the scaling factor is larger than in the 2D case. This flux, however, is a localized effect of the heterogeneity above the measurement profile. The enforced deviation at the surface causes the entire profile to be a worse representation of the observed dynamics. This results in an overall increase of the state uncertainty, which, in our case, is preferential because the uncertainty is generally underestimated.

Although fluxes between the measurement locations can be estimated, fig. 5.8 also exemplifies that the water content cannot, because the simplified estimated scaling field is not able to follow the true heterogeneity. Therefore, the EnKF in our application is only able to represent the true water content at the measurement locations. The estimated water content provides almost no information about the true water content in unobserved locations of the domain. Simultaneously, the horizontally averaged water content is a bad description of

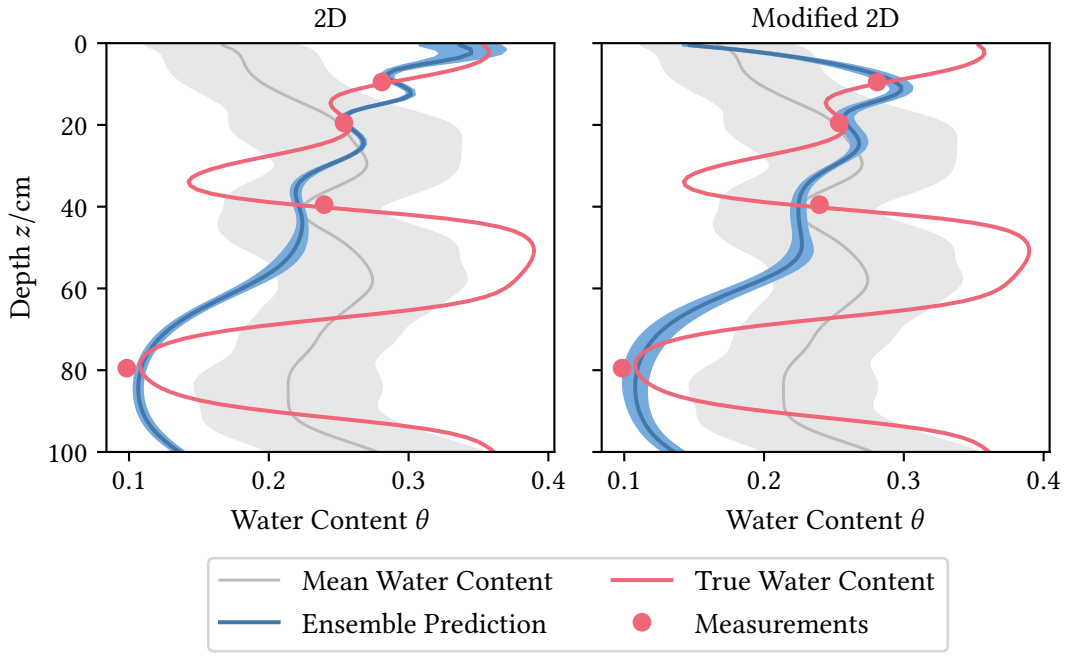

Figure 5.8: Soil hydraulic state for the regular and modified 2D case of the *HighVar* scenario with the worst prediction results at time $t = 29\,\mathrm{d}$. True water content and ensemble prediction are evaluated at the center of the domain, whereas the mean water content is the horizontal average. The shaded areas denote the respective standard deviation. The measurement uncertainty is roughly the width of the measurement markers.

the local water content and the associated measurements, and hence not represented in the ensemble either.

This demonstrates the interplay of model prescription and observations used in the data assimilation, as meaningful results can only be extracted from where the model is actually able to represent the desired information. If a soil is locally heterogeneous, then state forecasts are only valid where this heterogeneity is represented or estimated. Depending on the overall material properties and the heterogeneity itself, averaging data does not necessarily yield additional insights if localized information is requested. In case of the flux, we seek information averaged over the entire domain. As the dominating drivers for fluxes are gravity and the forcing at domain boundaries, it is crucial that the boundary fluxes at the measurement profile do not deviate strongly from the average boundary conditions. Especially evaporation, however, directly depends on the conductivity in the topmost part of the soil. Thus, material properties that strongly deviate from the average induce non-

representative fluxes there. If the single measurement profile is placed at such a location, a reasonable representation of the average fluxes in the domain requires adjustments of material properties or boundary conditions.

5.3.3 Conclusion

In this section and in the presented study we demonstrate that the soil hydraulic dynamics of a two-dimensional domain with small-scale heterogeneity and significant horizontal fluxes can be estimated with a one-dimensional domain and a reduced representation of small-scale heterogeneity. The accuracy of this estimation hinges on the exact configuration. Estimation errors become most obvious when comparing predicted against true fluxes, because the EnKF only operates on measurements of the water content in our case. However, this information is often difficult to obtain experimentally.

A large RMSE of the water content usually indicates a bad prediction for the flux and thus a low NSE. We find that the relative residual of the water content, where the absolute residual is divided by the ensemble spread, is a useful marker for structural model errors. If the prediction ensemble represented the uncertainty correctly, the relative residual would be <1 for approx. 68 % of the time, and <3 for approx. 99.7 % of the time, assuming the water content states would remain normally distributed. The latter is rarely the case, but a relative residual with values significantly larger than 3 should indicate considerable representation errors nonetheless. We used the synthetic truth to calculate the relative residual. But it can also be computed from the measurements, which yields the same quantity with an additional measurement error.

The investigated case with strong surface heterogeneity demonstrates that the measurement profile must be placed where the material properties of the soil near the surface do not deviate strongly from the mean material properties. This should be achievable in most experimental setups; in our example, the local length scale of the medium pore space was ten times lower than the average length scale, and this relation directly translates to the grain size distribution. Alternatively, local deviations above the measured soil profile can be taken into account in the data assimilation, as we achieved in the *Modified 2D* case. Even strong local deviations become less important with increasing depth, as the diffusive dynamics decrease the discrepancies emerging through the interplay of boundary conditions and errors in the estimation of local material properties.

6 Additional Effects of Unrepresented Topography

Real-world soil surfaces are rarely flat but feature a variety of shapes resulting primarily from erosion. With DORiE's ability to use unstructured grids, we can compute solutions of the Richards equation on domains with slanted or hill-like topographies. This gives new opportunities for investigating unrepresented model errors akin to the experiments conducted in chapter 5. In the absence of a preferred wind direction, precipitation and thus infiltration can be considered parallel to the acceleration of gravity. Therefore, a slanted surface will have a lower (normal) infiltration flux than a horizontal one by a factor of $\cos \nu$, where ν is the slope angle of the surface. Evaporation, on the other hand, is driven by the vapor pressure between the water phase at the surface and the surrounding air. Assuming that the air is a sufficiently large, well-mixed compartment, evaporation is independent of the topography and applies on any surface equally. This, however, can only be reasonably assumed on small spatial scales. In this chapter, I want to investigate the effects of model errors introduced through unrepresented topography of the domain by estimating soil hydraulic states and parameters of a domain with shaped surface using a model with flat surface in synthetic experiments.

Introducing a more complicated topography also calls for a discussion of associated processes DORiE does not consider. In the Richards regime, any infiltration flux into a soil can be supported by an appropriate adjustment of the matric head gradient. If the infiltration flux exceeds the saturated hydraulic conductivity, the matric head becomes positive, indicating that a force larger than the atmospheric pressure acts on the water phase to achieve infiltration at the boundary. Assuming that precipitation is the cause of infiltration, this force can only be supplied by gravity acting on the infiltrating water. Therefore, fluxes exceeding the saturated hydraulic conductivity typically imply water ponding on the surface, which in turn causes water runoff if the surface is tilted. As entrapped air often prohibits infiltration to saturate the soil matrix completely, runoff can be observed for lower precipitation fluxes as well. Its effect even on the scale of small catchments is considerable, as surface runoff was observed to contribute to total runoff by approx. 80 % for high intensity rainfall events in semi-arid regions (van Schaik, Schnabel, and Jetten 2008).

Nonetheless, surface runoff does not need to be considered in every scenario. In a hillslope experiment with sandy soil at the surface, Botto, Belluco, and Camporese (2018) did not observe surface runoff when applying irrigation fluxes of $1.63 \times 10^{-5}\,\mathrm{m\,s^{-1}}$ ($1408\,\mathrm{mm\,d^{-1}}$, some 16 % of the estimated saturated hydraulic conductivity) for several hours. Gevaert

et al. (2014) observed ponding and surface runoff on a sandy loam hillslope only after 22 h of continuous irrigation of $3.33 \times 10^{-6}\,\mathrm{m\,s^{-1}}$ ($287.7\,\mathrm{mm\,d^{-1}}$). The authors did not estimate the saturated hydraulic conductivity, however, and possible values ranged from $7.8 \times 10^{-6}\,\mathrm{m\,s^{-1}}$ to $1.4 \times 10^{-4}\,\mathrm{m\,s^{-1}}$, placing the irrigation flux in a range from 2 % to 42 % of saturated hydraulic conductivity.

In the study presented in chapter 5, the largest infiltration flux applied on a *Sandy Loam* domain was $4 \times 10^{-7}\,\mathrm{m\,s^{-1}}$ over a time span of 2 d, some 3.3 % of the saturated hydraulic conductivity of the homogeneous reference material, cf. table 1.1. However, the scaling field ξ modifies the local saturated hydraulic conductivity by a factor of ξ^2. In the *HighVar* scenario, the infiltration flux can therefore reach the saturated hydraulic conductivity at locations with very low scaling factors. In the *MedVar* scenario with $\sigma_\xi^2 = 0.2$, statistically only approx. 2.5 % of the domain has a lower saturated hydraulic conductivity than

$$K_0 = K_0^* \cdot \exp^2\left(2\sigma_\xi - \sigma_\xi^2\right) = 1.34 \times 10^{-6}\,\mathrm{m\,s^{-1}}, \tag{6.1}$$

of which the strongest applied infiltration flux was approx. 30 %. One could thus expect ponding at these locations in the *MedVar* scenario, but the water could then infiltrate in surrounding areas, as the correlation lengths of the random field are relatively short. It is therefore reasonable to assume that the parameterization and boundary conditions of the *MedVar* scenario could be applied to a tilted surface as well, without considering surface runoff. This also relieves us from incorporating processes resulting from runoff, like water ponding in surface depressions and erosion, which modifies the domain over time.

6.1 Synthetic Experiments

Similar to the study presented in chapter 5, I generated a synthetic truth and synthetic measurements from it by applying the Richards solver of DORiE on a two-dimensional domain with fully resolved heterogeneity. I then estimated the soil hydraulic state and parameters, and a reduced representation of the heterogeneity using a pseudo one-dimensional model. In this case, the synthetic truth involved a more complicated surface shape. Thus, in addition to the heterogeneity at full resolution, the estimation model also did not represent the surface shape. Both features led to inherently two-dimensional flow, which could not be represented in the estimation model. Additionally, the surface itself has implications on the boundary conditions. The measurement setup remained the same, with four measurements being placed in depths of 9.5 cm, 19.5 cm, 39.5 cm and 79.5 cm below the surface. The settings for the data assimilation and the reduced representation of heterogeneity by only five estimated Miller scaling factors (at the measurement location and below the surface) with interpolation were directly adopted for the experiments in this chapter.

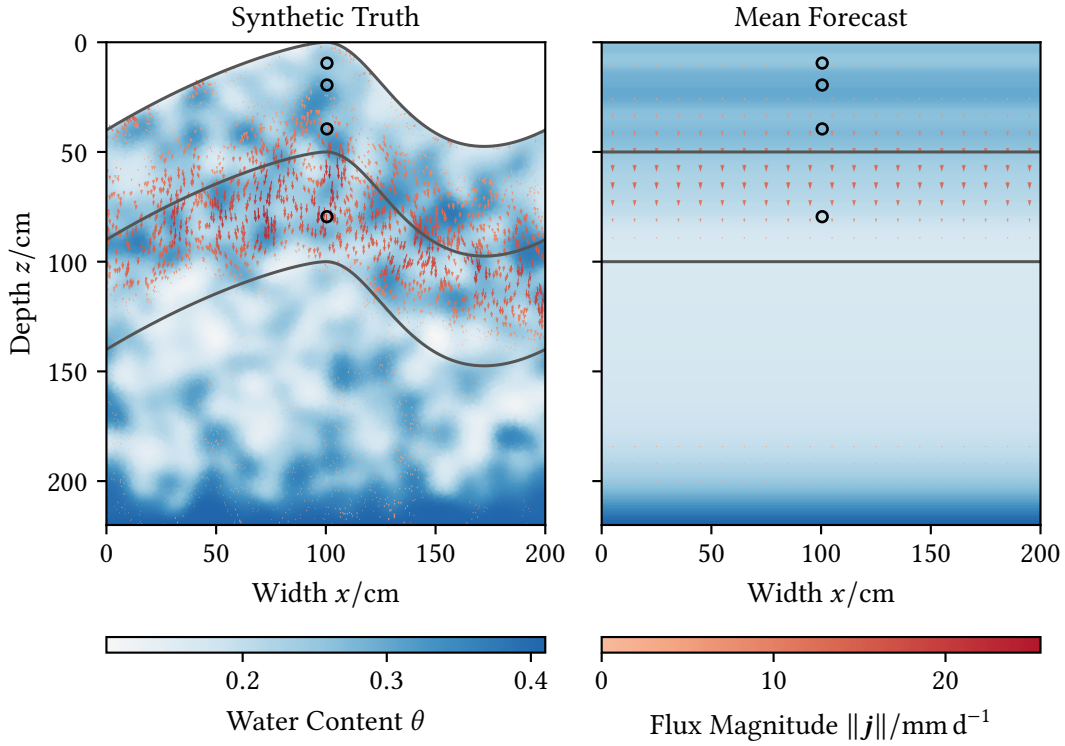

Figure 6.1: Hydraulic state and flux of the two-dimensional synthetic truth (left) and the mean ensemble forecast (right) at time $t = 43$ d for one case of the *MedVarHill* scenario. Gray lines indicate the evaluation positions of the perpendicular flux at a depth of 0 cm, 50 cm and 100 cm from the surface. Measurement positions are indicated by black circles. Arrows indicate flux direction and magnitude. The color bars apply to both panels.

6.1.1 Model Setup

The simulated domain had overall extensions of 2 m×2.2 m, with the highest elevation in the center (Figure 6.1, left panel). Towards the left, I modeled a continuous slope reducing the elevation to 1.8 m at the left boundary. Towards the right, I chose a steeper slope, followed by a depression and a slight upward slope towards the right boundary. The elevation difference of the domain was nearly 50 cm. I chose the highest point of the domain to have a depth of $z = 0$ m, meaning that other parts of the surface were located at positive depths.

I applied the same boundary conditions as in the previous chapter, cf. section 5.2.1. Only the outflow boundary condition, which modeled evaporation, had to be adjusted because of the elevation difference. Applying a constant matric potential ψ_m on a tilted surface would

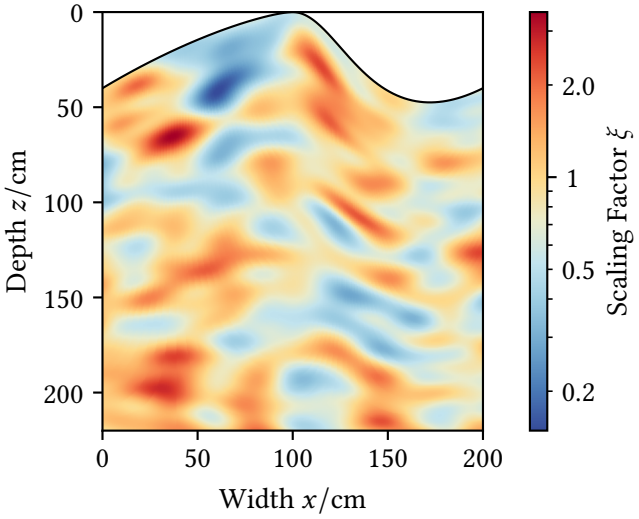

Figure 6.2: Instance of a surface-distorted Miller scaling field used in the scenario *Shifted*. This is the result of a modified evaluation of a scaling field with axially parallel anisotropy. The applied shift causes the heterogeneous pattern to follow the surface in low depths, and recovers the original anisotropy towards the bottom of the domain.

induce a gradient parallel to the boundary due to changes in the gravitational potential ψ_g. However, the atmospheric forcing induces a (total) fluid potential ψ_w which, in the soil, is the sum of the matric potential and the gravitational potential (Roth 2012),

$$\psi_w = \psi_m + \psi_g = h_m \rho g - z \rho g. \tag{6.2}$$

Assuming a constant fluid potential in the atmosphere over the entire domain, this yields a boundary condition potential ψ_m which is in hydraulic equilibrium with itself, meaning that the boundary condition only induces fluxes perpendicular to the boundary. With the original boundary condition value u_D to be applied at the surface, I therefore chose a modified boundary condition of

$$u'_D(z) = u_D + z. \tag{6.3}$$

The infiltration boundary conditions were adjusted to the surface as well. I envisaged the cause of the infiltrating water being precipitation or irrigation. Assuming negligible wind, these fluxes apply parallel to the gravitational acceleration and are measured with respect to a unit area. Neumann boundary condition fluxes, however, are applied perpendicular to the boundary. If a surface is tilted by an angle ν, one has to multiply the irrigation flux by a factor of $\Delta = \cos \nu$ to ensure that the amount of infiltrated water per unit area remains constant over the domain. Alternatively, one can compute this factor with the scalar product of the boundary unit normal vector and the unit normal vector in direction of gravitational acceleration, $\Delta = |\hat{\boldsymbol{n}}_F \cdot \hat{\boldsymbol{g}}|$. DORiE supports this adjustment via a simple key in the Neumann boundary condition definition.

Like in the last chapter, I chose the only difference between all cases to be the Miller scaling

field used when simulating the synthetic truth. In a first step, I generated new fields covering the now extended 2 m × 2.2 m domain, yielding five cases of the *MedVarHill* scenario. DORiE only accepts rectangular scaling fields and applies them on the field spanned by the domain maxima by default, which simply disregards field values that lie outside the mesh. In a separate scenario, I created scaling fields with locally varying anisotropy, as opposed to the anisotropy characterized by axially parallel, vertical and horizontal correlation lengths used until now. To that end, I first generated a high-resolution scaling field with axially parallel anisotropy on a 3 m × 3 m domain. I then extracted a portion of the size of the simulated domain while applying distortion. With the original scaling field ξ, the values at width x and depth z were given by

$$\xi'(x, z) = \xi\left(x, z'(x, z)\right), \tag{6.4}$$

with the shifted evaluation depth

$$z'(x, z) = \underbrace{\frac{2.2\,\mathrm{m} - z}{2.2\,\mathrm{m}}}_{\text{decrease with depth}} \cdot \overbrace{\left(S(x) - 0.2\,\mathrm{m}\right)}^{\text{shift according to surface}} + \underbrace{0.5\,\mathrm{m}}_{\text{evaluation offset}}, \tag{6.5}$$

where $S(x)$ is the depth of the surface at width x. This vertical shift achieves that the local anisotropy of the field is axially parallel to the surface tangent and normal. The shift is linearly reduced with increasing depth, reaching a value of zero at the lower domain boundary. The additional offset is used to avoid evaluating the original field ξ outside its bounds. This procedure produced a field with anisotropic statistical properties. For the original scaling field, I chose a larger horizontal correlation length of $\ell_h = 20$ cm to achieve a layer-like pattern while avoiding the situation where the medium appears locally layered, as in the *Horizontal* scenario of chapter 5. At the same time, this choice of the horizontal correlation length also avoided dominantly vertical features in the shifted field. A resulting field used to compute the synthetic truth is displayed in fig. 6.2, and the scenario was named *ShiftedHill*.

As the FV scheme of DORiE cannot be used on unstructured grids, I computed the two-dimensional synthetic truth and the synthetic measurements using the DG scheme with an unstructured grid built with the GMSH application (Geuzaine and Remacle 2009). Its simplex grid elements had an average edge length of 1 cm. The grid for the ensemble filter forward model was extended to span the domain of 2 m × 2.2 m with a vertical grid resolution of $h_T = 1$ cm. This model used the FV scheme for reducing the computational cost and to avoid failing forecasts after state updates, as discussed in section 4.3. In this series of experiments, I did not compute the one-dimensional control cases, but instead computed a single two-dimensional *Hill* control case without heterogeneity to investigate the effects of unrepresented topography alone and compare it to the cases with heterogeneity.

For the initial condition of the synthetic truth I dropped the 30 d spin-up phase implemented by Bauser et al. (2020) and instead computed a stationary solution from an infiltration flux of $5 \times 10^{-9}\,\mathrm{m\,s^{-1}}$ ($0.432\,\mathrm{mm\,d^{-1}}$) into the domain. This feature was not yet implemented in DORiE at the time the experiments discussed in chapter 5 were conducted and significantly reduced the computational cost for the initial condition. While the solution of the stationary problem is not equivalent to that of the spin-up simulation, it achieves a similar initial condition: the domain is not in hydrostatic equilibrium, but the current fluxes are negligible with respect to the boundary conditions applied during the assimilation and verification phases of the experiments. The initial guess for the water content in the data assimilation was again achieved by interpolating synthetic measurements of the initial condition and distorting each state with localized values drawn from an unbiased normal distribution (Section 5.2.2).

6.1.2 Evaluation

For evaluating the results I used the quantities and descriptors AHFR, RMSE_θ, ΔJ, and NSE_j, cf. section 5.2.3 for the definitions. The shaped surface present in the synthetic truth, however, required a change in the evaluation of the true flux used in the computation of ΔJ and NSE_j. The one-dimensional model used in the estimation does not account for the surface shape. Thus, fluxes will be vastly different when comparing them at the same depth z because of different distances to the surface, where infiltration and evaporation applies. For a sensible comparison, I evaluated the true fluxes at constant depths from the surface (Figure 6.1) by averaging the flux component normal to the surface $j_\perp$ over the surface curve length s. I denote this averaged normal flux by $\langle j_\perp \rangle_s$, changing the definition of the relative cumulative flux error to

$$\Delta J = \frac{\sum_t^{N_t} \left[\overline{j_{\perp,t}} - \left\langle j^*_{\perp,t} \right\rangle_s \right]}{\sum_t^{N_t} \left\langle j^*_{\perp,t} \right\rangle_s}, \tag{6.6}$$

and that of the flux NSE to

$$\mathrm{NSE}_j = 1 - \frac{\sum_t^{N_t} \left[\overline{j_{\perp,t}} - \left\langle j^*_{\perp,t} \right\rangle_s \right]^2}{\sum_t^{N_t} \left[\left\langle j^*_{\perp,t} \right\rangle_s - \sum_t^{N_t} \left\langle j^*_{\perp,t} \right\rangle_s / N_t \right]^2}, \tag{6.7}$$

respectively. In the one-dimensional estimation model, the flux perpendicular to the surface is equivalent to the vertical flux, $j_\perp \equiv j_z$.

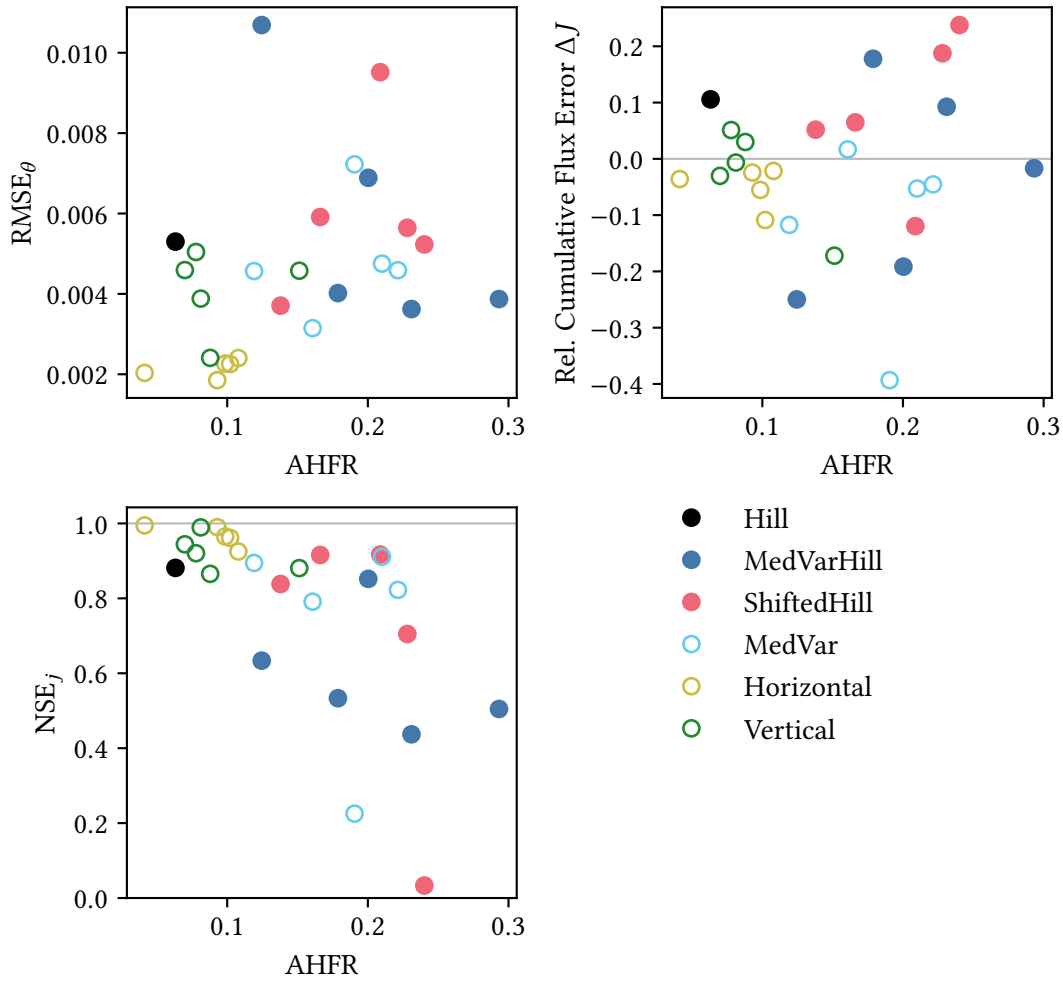

Figure 6.3: Summary measures for experiments conducted using unrepresented heterogeneity and topography. The water content root mean squared error RMSE$_\theta$, relative cumulative flux error ΔJ, and flux Nash-Sutcliffe efficiency NSE$_j$ are plotted against the average horizontal flux ratio AHFR, which measures the heterogeneity of the flow field. Colors indicate the scenario. Filled markers denote the cases of this study, and outlined markers denote three scenarios from chapter 5 for comparison. Gray lines highlight the ideal values $\Delta J = 0$ and NSE$_j = 1$.

Table 6.1: Overview and results of experiments conducted using unrepresented heterogeneity and topography. The center column group lists the properties of the generated Miller scaling fields. (★) denotes the properties of the original scaling field before shifting was applied. The right column group contains the averaged results of the experiments. The overline denotes the mean calculated from the five cases run for each combination. (†) indicates that no mean was taken as only a single value was given. Results for *MedVar*, *Horizontal*, and *Vertical* scenarios are taken from chapter 5 and listed for comparison.

Scenario	Miller Scaling Field			$\overline{\text{AHFR}}$	$\overline{\text{RMSE}_\theta}$	$\overline{\Delta J}$	$\overline{\text{NSE}_j}$
	$\sigma^2_{\tilde{\xi}}$	ℓ_h/cm	ℓ_v/cm				
Hill	0.0	/	/	$0.06^\dagger$	$0.0053^\dagger$	$0.11^\dagger$	$0.88^\dagger$
MedVarHill	0.2	10	10	0.21	0.0058	−0.04	0.59
ShiftedHill	$0.2^\star$	$20^\star$	$10^\star$	0.20	0.0060	0.08	0.68
MedVar (2D)	0.2	10	10	0.18	0.0049	−0.13	0.73
Horizontal (2D)	0.1	30	10	0.09	0.0022	−0.05	0.97
Vertical (2D)	0.1	10	30	0.09	0.0041	−0.06	0.92

6.2 Results and Discussion

The cases from the *MedVarHill* and *ShiftedHill* scenarios are best compared to the *MedVar* scenario of chapter 5, which featured the same scaling field variance. We find that the hill-shaped surface in the synthetic truth increases the horizontal flux component and thus the overall AHFR, with mean values of 0.2 and 0.21 for *ShiftedHill* and *MedVarHill* against values of 0.18 for *MedVar* and 0.25 for *HighVar* (Table 6.1). The RMSE of the water content is increased and the NSE of the flux is decreased, indicating a worse predictive performance of the estimated model. Similar to the AHFR, the gray mean NSE_j for both new scenarios lies roughly in between the values computed for *MedVar* and *HighVar*. Nonetheless, the mean relative cumulative flux error ΔJ lies comparably close to zero. The tendency in the measures is similar to the ones reported in chapter 5 (Figure 6.3). The RMSE of the water content generally increases with increasing AHFR, while the NSE of the flux decrease. The conducted cases show no clear tendency in the relative cumulative flux error, but its spread of values for a single scenario is large with a range of approx. 30 % and 40 % in the *ShiftedHill* and *MedVarHill* scenarios, respectively. The single case with homogeneous mean in the *Hill* scenario has a low AHFR of 0.06 resulting only from the horizontal flux components of the infiltration and evaporation boundary conditions. With an RMSE_θ of 0.0053, a relative cumulative flux error $\Delta J = 11$ %, and a flux NSE of 0.88 it yields nearly equivalent results to the other scenarios of this section with heterogeneous domain.

Although $\overline{\text{AHFR}}$ and $\overline{\text{RMSE}_\theta}$ are similar, the estimation in the *ShiftedHill* scenario yields a significantly better flux NSE than in the *MedVarHill* scenario, with $\text{NSE}_j = 0.68$ against a value of 0.59. This can be explained with the heterogeneous soil structure following the surface shape. In the upper part of the domain, the heterogeneity forms layers effectively parallel to the surface. Fine-grained layers divert strong infiltration fronts in directions parallel to the surface, whereas coarse-grained layers block evaporation from center regions of the hill. Both phenomena reduce effects from the surface shape on the measurements and therefore make the one-dimensional model a better representation of the observed dynamics. One could expect that the above effects would be inverted if the measurements were placed below the hill slopes, or especially below the depression towards the right boundary of the domain, where infiltration fluxes from below the slopes are expected to be converging.

With these results, I conclude that data assimilation of a two-dimensional, heterogeneous domain with complicated topography is possible with a one-dimensional model and a reduced representation of heterogeneity. The errors in the estimation of water content and fluxes increase compared to scenarios which only feature unrepresented heterogeneity, but in a similar way as if only the heterogeneity was increased. In the following subsections, I will investigate how to improve the flux estimation in scenarios with unrepresented topography and heterogeneity, and attempt to combine results from all experiments presented in chapter 5 and this chapter.

6.2.1 Modeling of Evaporation

The boundary condition applied in the one-dimensional estimation model was chosen to represent the boundary condition directly above the measurement profile of the synthetic truth. While the adjusted infiltration boundary conditions are equivalent, the evaporation boundary condition of the synthetic truth is adjusted to account for the surface shape and decreases with increasing depth of the surface. Therefore, the estimation model applies a stronger outflow boundary condition through a lower matric head and generally overestimates evaporation. This has a profound effect on the estimation of the flux, resulting in a flux NSE of 0.88 and a cumulative flux error ΔJ of 11 % in the *Hill* scenario, where the synthetic truth was a homogeneous domain.

The average depth of the topography depicted in figs. 6.1 and 6.2 is $z = \langle S(x) \rangle \approx 24.24\,\text{cm}$. To better resemble the mean evaporation from the domain modeled in the synthetic truth, I adjusted the outflow boundary condition of the estimation model by reducing the initial value from $h_m = 0\,\text{m}$ to $h_m = 24.24\,\text{cm}$ while maintaining the temporal increase, and repeated a single experimental case of the *MedVarHill* scenario. This particular case is characterized by an AHFR of 0.12 and the estimation results were $\text{RMSE}_\theta = 0.011$, $\Delta J = -25\,\%$, and $\text{NSE}_j = 0.63$ without modified boundary condition.

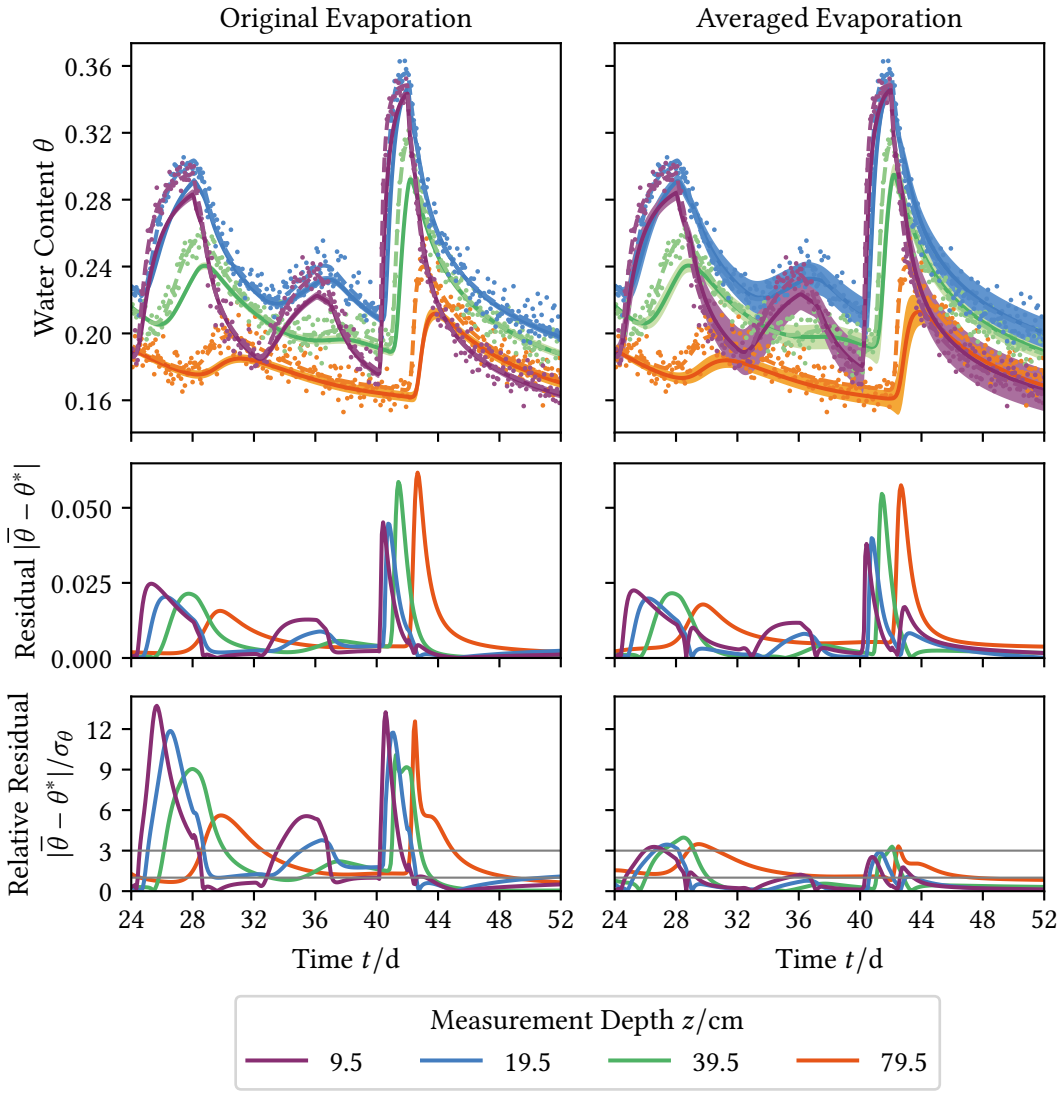

Figure 6.4: Water content predictions at the four measurement depths for the original and modified boundary conditions of one case of the *Hill* scenario. The top panels show the mean of the ensemble forecast (dark lines) with respective standard deviation (shaded areas), the synthetic truth (dashed lines), and the noisy measurements (markers). The center panels display the absolute residuals between the ensemble mean $\overline{\theta}$ and the synthetic truth θ^*, and the lower panels show the same values divided by the respective standard deviation σ_θ computed from the ensemble at this location. The gray lines indicate the σ_θ- and $3\sigma_\theta$-uncertainty ranges.

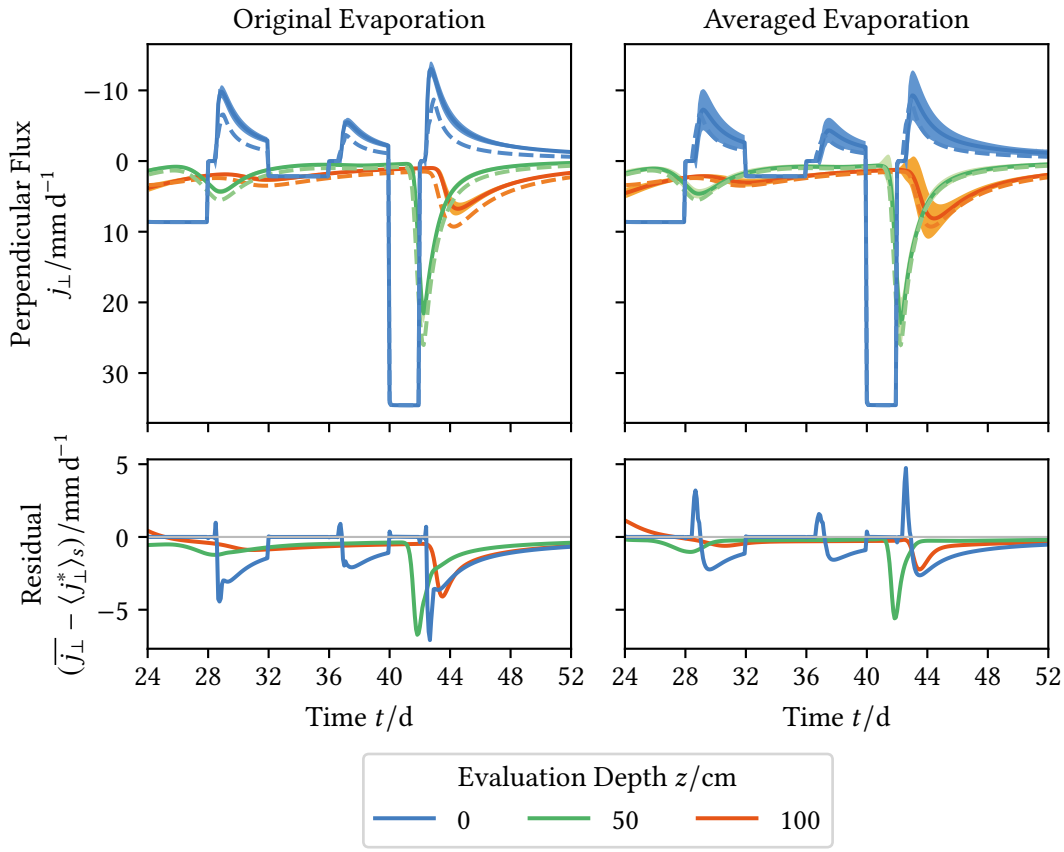

Figure 6.5: Flux predictions at three evaluation depths for the original and modified boundary conditions of one case of the *Hill* scenario. The top panels show the mean of the ensemble forecast (dark lines) with respective standard deviation (shaded areas) and the synthetic truth (dashed lines). The perpendicular flux $j_\perp$ is the flux component perpendicular to the respective evaluation face displayed in fig. 6.1. Note that the vertical axis is inverted: a positive perpendicular flux component on a flat surface indicates a downward flux. The lower panels display the residual between the ensemble mean $\overline{j_\perp}$ and the synthetic truth $\langle j_\perp^* \rangle_s$.

With the reduced outflow boundary condition the results improved to $\mathrm{RMSE}_\theta = 0.010$, $\Delta J = -10.6\,\%$, and $\mathrm{NSE}_j = 0.90$, indicating a slight improvement in the forecast of the water content, but a significant improvement one for the flux prediction. The most notable change was in the ensemble spread σ_θ, which increased by a factor of > 4 (Figure 6.4). This meant that the relative residual dropped into the $3\sigma_\theta$-range, indicating that the model correctly represented model errors in the state uncertainty. The increase in the ensemble spread was achieved by the state inflation. Compared to the original evaporation boundary condition, the modified, averaged boundary condition provides a worse description of the evaporation right above the measurement profile. The increased matric head applied means that the evaporation flux in the estimation model starts later than in the synthetic truth (Figure 6.5). This increases the deviation between measurements and estimated state in the topmost parts of the soil. The EnKF reacts to this discrepancy using the inflation. Increasing the ensemble spread, this lets the filter adjust the state to the measurement values which cannot be explained by the dynamics of the estimation model. As the modified outflow boundary condition is a better description of the average evaporation from the true domain, the flux residual was significantly reduced.

These results emphasize the importance of the inflation method used in the EnKF. It is required not only to adjust the estimated state to measurements that cannot be explained by the modeled dynamics, but also for a sensible estimation of the state uncertainty, which is encoded in the ensemble spread. Bauser et al. (2018) already stressed and demonstrated that a specialized inflation method is key for achieving valuable estimation results from models with unrepresented model errors. Their inflation method, which is also applied in the EnKF used in all experiments of this thesis, features a spatio-temporal adaptation based on deviations between measurements and the mean ensemble state. The authors point out that this is an important feature because for data assimilation in soil hydrology, inflation methods must be able to locally and rapidly increase and decrease ensemble inflation. Apparently, temporally short increases of inflation are sufficient to resolve most discrepancies between estimation model and measurements due to unrepresented heterogeneity and topography. Because the applied boundary conditions are the same as in the synthetic truth, errors in the water content are quickly accounted for by adjusting soil hydraulic parameters, local scaling factors, and the water content state itself. The inflation then quickly decreases, thus underestimating the model error. The modified boundary condition, on the other hand, produces a temporally extended deviation, which induces longer periods of inflation during the assimilation phase. This results in a broader ensemble at the end of that phase and consequently also in the verification phase.

One could thus interpret the improved representation of uncertainty as artifact of an artificially increased inflation. With respect to the observations, the modified boundary condition increased the discrepancy between synthetic truth and estimation model. This yielded a stronger and more continuous signal to increase inflation, which ultimately improved the estimation. These findings emphasize the need for the right amount and

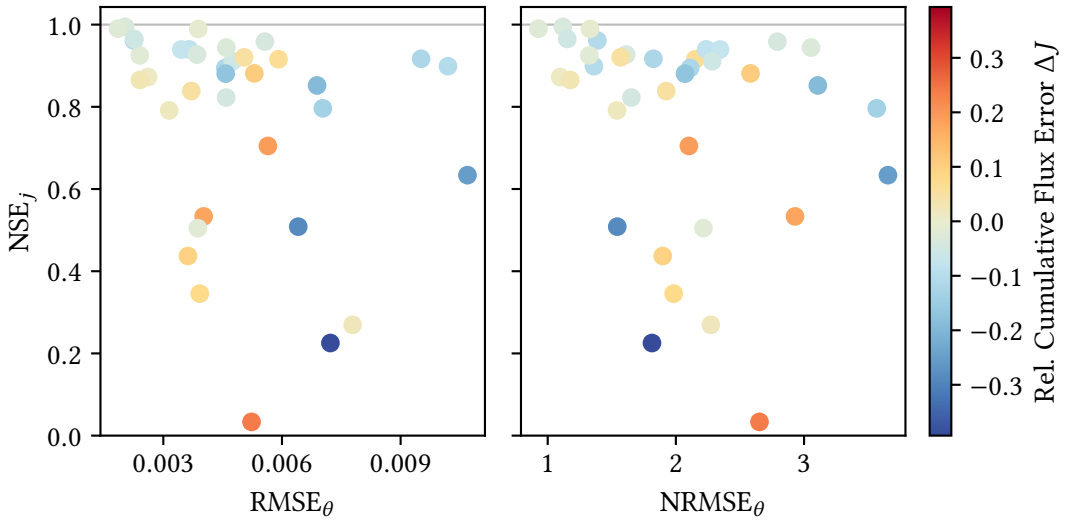

Figure 6.6: Flux Nash-Sutcliffe efficiency NSE_j plotted against the water content root mean squared error RMSE_θ (left) and normalized root mean squared error NRMSE_θ (right) for 2D cases of all experiments discussed in chapters 5 and 6, except for the single case with negative NSE_j discussed in section 5.3.2. The gray line indicates the ideal value of $\mathrm{NSE}_j = 1.0$.

duration of inflation expressed by Bauser et al. (2018), which depends on the model and the application. In our case, a temporally less adaptive inflation method generally might have yielded better results, especially in terms of estimated model uncertainty. Analogous to the findings in section 5.3.2, we can also conclude again that the correct representation of the boundary conditions over the entire domain improve the estimation result.

6.2.2 Overall Estimation and Forecast Accuracy

The nature of the synthetic experiments discussed and conducted within this thesis allowed for comparisons of different quantities between the synthetic truth and the ensemble states of the simplified model used for assimilation and forecast. This enabled discussions on the performance of the estimation and the quality of the forecast. In the analysis, I focused on the accuracy of the flux prediction because the flux is a better descriptor of the overall dynamics than the water content at the measurement locations. But in a real world scenario, the information on the true flux will hardly be available. I therefore studied the quantities retrievable from the measurements alone to find relationships between them and the flux forecast accuracy.

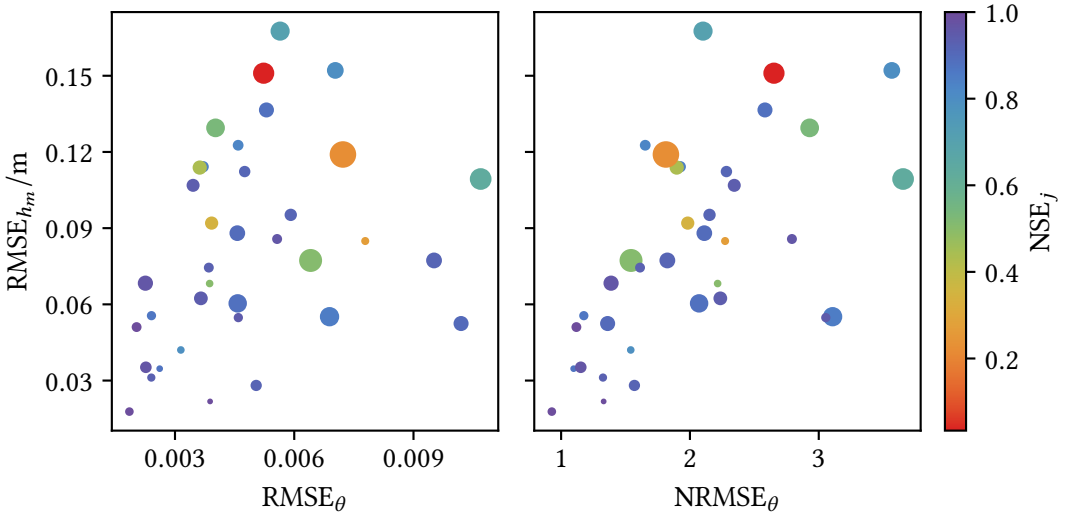

Figure 6.7: Matric head root mean squared error RMSE_{h_m} plotted against the water content root mean squared error RMSE_θ (left) and normalized root mean squared error NRMSE_θ (right) for 2D cases of all experiments discussed in chapters 5 and 6, except for the single case with negative flux Nash-Sutcliffe efficiency NSE_j discussed in section 5.3.2. Marker colors indicate NSE_j and marker sizes denote the respective absolute value of the relative cumulative flux error ΔJ ranging from 0.7 % to 39.3 %.

The ensemble spread in a dimension of the EnKF state vector gives the estimation accuracy of the state in this particular dimension. Its value can be interpreted as the (inverse) confidence in the currently estimated state and thus as the estimated uncertainty in this state considering the estimation model, the measurements, and possibly unrepresented model errors. As speculated by Bauser et al. (2020) and in section 5.3.3, an error between estimated state and measurements that has been normalized based on the ensemble spread could be a useful indicator for the estimation accuracy. I therefore define the normalized root mean squared error (NRMSE) of the water content as

$$\mathrm{NRMSE}_\theta := \sqrt{\frac{1}{N_t N_m} \sum_{t}^{N_t} \sum_{m}^{N_m} \left[\frac{\theta^*_{t,m} - \overline{\theta_{t,m}}}{\sigma_{\theta,t,m}} \right]^2}, \tag{6.8}$$

where $\sigma_{\theta,t,m}$ is the standard deviation of the water content state at time t and measurement location m in the state ensemble.

If the ensemble spread was large enough to accurately represent the deviations between the ensemble mean and the truth, we would expect that $\mathrm{NRMSE}_\theta \leq 1$. However, such a value is

only reached in a single case from all discussed experiments, which affirms the preceding conclusion that the state ensemble generally underestimates the uncertainty (Figure 6.6). We find that RMSE_θ and NRMSE_θ alone are not able to indicate a tendency for the flux NSE and the cumulative flux error. While experimental cases with very low values of RMSE and NRMSE also have a flux NSE close to one, the NSE_j may range from 1 to nearly 0 for any observed RMSE value. Still, a low NSE_j also implies a low cumulative flux error $|\Delta J|$.

While fluxes inside the soil are difficult to observe, measurements of the matric potential at specific locations are possible, e.g. with tensiometers. Therefore, it is experimentally feasible to use water content measurements for data assimilation, as done in the experiments discussed in this thesis, and then verify the results by additionally using measurements of the matric potential. Analogous to the water content RMSE, I define the matric head RMSE for our setup as

$$\mathrm{RMSE}_{h_m} \coloneqq \sqrt{\frac{1}{N_t N_m} \sum_{t}^{N_t} \sum_{m}^{N_m} \left[(h_m^*)_{t,m} - \overline{(h_m)_{t,m}} \right]^2}, \tag{6.9}$$

where $(h_m^*)_{t,m}$ and $\overline{(h_m)_{t,m}}$ indicate the true and mean estimated matric head at time t and measurement location m, with said locations being the same as for the water content. Both values are computed by inserting the respective water content value in the inverse of the retention curve, eq. (1.13), using the true or estimated material properties and scaling factors at location m.

For any RMSE_θ, the resulting RMSE_{h_m} can take a wide range of values (Figure 6.7). Nonetheless, these values are relatively small, as all considered experiments yield $\mathrm{RMSE}_{h_m} < 20\,\mathrm{cm}$. Combining the results for NRMSE_θ, RMSE_{h_m}, NSE_j, and $|\Delta J|$ indeed reveals a clearer connection: The accuracy in the flux prediction is generally high if both NRMSE_θ and RMSE_{h_m} are low. For the experiments discussed in this thesis, we find that $\mathrm{NSE}_j > 0.5$ for all cases with $\mathrm{NRMSE}_\theta < 2$ and $\mathrm{RMSE}_{h_m} < 9\,\mathrm{cm}$. Nonetheless, high flux NSE values are also possible for higher RMSE values. While these results only include the *MedVarHill* and *ShiftedHill* scenarios with original evaporation boundary condition, we can assume that repeating the cases with the modified boundary condition matches this relation: through the increased inflation and ensemble spread, the NRMSE_θ will be significantly decreased, while the NSE_j will be increased.

I conclude that the RMSE of the assimilated variable alone is an insufficient descriptor of the estimation accuracy and additional quantities have to be measured and compared. This can be explained with the basic principle of the EnKF, which aims to minimize discrepancies between the ensemble of estimated states and measurements based on correlations. As the ensemble spread and thus the model uncertainty is generally underestimated, the NRMSE alone cannot give more insight into the accuracy either. However, my analysis indicates

that the results based on water content measurements can be combined with observations of other quantities to yield a better estimation of the accuracy and predictive capabilities of the EnKF results.

Summary

The physics of soil water flow describes the essential hydrological process for many ecosystem services at a key terrestrial interface. The complicated multi-scale architecture of soils and their highly non-linear fluid dynamics require robust numerical solvers and a good knowledge of material properties to compute reliable solutions. My research focused on the development, implementation, and analysis of tools for numerically solving the Richards equation and for estimating soil hydraulic states and parameters. In this thesis, I analyzed the DUNE Operated Richards Equation Solving Environment (DORiE; Riedel et al. 2020a) for computing soil water flow with the finite volume (FV) and the discontinuous Galerkin (DG) method, and applied it in synthetic experiments to research the accuracy and robustness of a data assimilation method when it encounters unrepresented model errors.

Using DORiE, I investigated the accuracy and performance of FV and DG discretization schemes for solving the Richards equation. DG schemes generally achieve a much higher accuracy than the FV method. They can be applied on unstructured grids of various shapes and also support local adaptive grid refinement (h-refinement), which is implemented in DORiE. This flexibility increases their computational cost and thus reduces their efficiency, especially in applications where accuracy is less important. In line with the related literature, I find that the DG discretization is not generally advantageous to the FV discretization of the Richards equation. The DG scheme is preferred when highly accurate solutions or solutions on unstructured grids with possibly complicated domain shapes are required. If the domain can be approximated by a regular grid and solutions are not required to be highly accurate, the FV method is favorable. Like similar studies, my benchmarks also indicate that h-refinement might bridge this gap, given that its algorithm is improved and optimized. As presented, DORiE implements a versatile Richards solver for various applications, including a DG discretization with h-refinement as unique feature among comparable software packages. Future work on DORiE could focus on improving the numerics to increase efficiency and broaden use cases of the DG method, e.g. by implementing p-refinement. This would make DORiE more suitable for the integration in other frameworks for inverse modeling, data assimilation, or large-scale hydrology.

As measurements of soil hydraulic states are typically noisy and scarce, robust methods for combining observations and model forecasts are needed for an accurate estimation of soil water flow and material properties. With several associated processes found in soils, models for soil water flow typically entail a considerable representation error. This is also the case

for the process of soil water movement itself: any model with a limited resolution is unable to represent the complete multi-scale, heterogeneous soil architecture. In a real-world scenario, this structure can typically only be estimated.

I therefore investigated the robustness and accuracy of an ensemble Kalman filter (EnKF; Evensen 2003) with specializations for soil hydrology using synthetic experiments. Some findings of this research on the effects of unrepresented small-scale heterogeneity were already published by Bauser et al. (2020). I revisited and further analyzed these results, and I expanded the research effort with experiments including unrepresented topography as additional model error. The main setup of all experiments was a two-dimensional, heterogeneous domain containing a one-dimensional soil profile, in which four synthetic measurements were conducted. This synthetic truth was then assimilated using a one-dimensional model with reduced representation of heterogeneity. The combined results show that the specialized EnKF is capable of estimating soil hydraulic states, water fluxes, and material properties even if topography and small-scale heterogeneity are not represented in the assimilation model. The estimation yields effective material properties which do not necessarily resemble the true properties but account for the errors in the representation. However, further analysis and modification of two experimental cases demonstrated that an accurate prediction of fluxes in the estimated domain requires boundary conditions applied in the estimation model to approximate the boundary conditions throughout this domain. In the given setup, forecast fluxes are sensitive to errors in the boundary condition representation, but the EnKF is able to compensate discrepancies between the assimilated soil hydraulic state and the boundary conditions by adapting the effective material properties. Using the presented methods, a one-dimensional model is capable of representing the essential dynamics of a two-dimensional, heterogeneous flow field with complicated topography.

The accurate representation of uncertainty remains an open question. While the EnKF was able to assimilate the observed dynamics in the majority of experimental cases discussed in this thesis, it usually underestimated the state uncertainty. This is a problem for real-world applications, where estimation results are difficult to verify and a low uncertainty in the filter state would lead to a high confidence in said results. One cause of the low uncertainty is the low influence of inflation in the experimental cases. The deviation from unrepresented heterogeneity and topography could usually be resolved too quickly to demand a continuous inflation. While the inflation method could be improved to better detect such errors, a possibly more viable option for increasing the uncertainty is to add an explicit model error. As such an error is specific to the application and the model used, it was set to zero in all discussed experiments. However, its estimation is the target of ongoing research in soil hydrology, and recent studies suggest that it can be estimated with purely data-driven methods. My analysis further indicates that additional observables—primarily the matric potential—can be used to verify the results if they have not been used for the data assimilation itself.

Acknowledgments

A significant part of my research was funded by Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg (Az 33-7533.-30-20/6/2).

Graphics in this thesis use color schemes based on the exceptional work of Paul Tol,[1] with the intention of maximizing accessibility for persons with color vision deficiency.

Countless people in my private and professional environment enabled me to conclude this thesis, and I am grateful to each and every single one of them. Being well aware that I will miss several important figures, I explicitly wish to thank:
Kurt for being a scientific and personal mentor during my time in Heidelberg,
Jule, Benjamin, Harald, Yunus, and all members of the TS-CCEES working group for the great, more or less scientific endeavors in and especially around the institute,
Hannes for his guidance and our prolific collaboration,
Santiago for the invaluable work he invested in DORiE,
Rebecca, Stefanie, and Markus for always supporting and never once doubting me,
Kilian for never growing tired of indulging in our numerous mutual interests, and
Daniela for her seemingly endless care, affection, and love.

1. See https://personal.sron.nl/~pault/.

Publications by the Author

Publications used in this thesis:

Bauser, H. H., Riedel, L., Berg, D., Troch, P. A., and Roth, K. 2020. "Challenges with effective representations of heterogeneity in soil hydrology based on local water content measurements." *Vadose Zone Journal* 19 (1): e20040. https://doi.org/10.1002/vzj2.20040.

Riedel, L., Ospina De Los Ríos, S., Häfner, D., and Klein, O. 2020a. "DORiE: A discontinuous Galerkin solver for soil water flow and passive solute transport based on DUNE." *Journal of Open Source Software* 5 (52): 2313. https://doi.org/10.21105/joss.02313.

Publications not referenced in this thesis:

Riedel, L., Herdeanu, B., Mack, H., Sevinchan, Y., and Weninger, J. 2020b. "Utopia: A Comprehensive and Collaborative Modeling Framework for Complex and Evolving Systems." *Journal of Open Source Software* 5, no. 53 (2020): 2165. https://doi.org/10.21105/joss.02165.

Sevinchan, Y., Herdeanu, B., Mack, H., Riedel, L., and Roth, K. 2020. "Boosting Group-Level Synergies by Using a Shared Modeling Framework." In *Computational Science – ICCS 2020,* edited by V. V. Krzhizhanovskaya, G. Závodszky, M. H. Lees, J. J. Dongarra, P. M. A. Sloot, S. Brissos, and J. Teixeira, 12143:442–456. Lecture Notes in Computer Science. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-030-50436-6_32.

Bibliography

Alexander, R. 1977. "Diagonally Implicit Runge-Kutta Methods for Stiff O.D.E.'s." *SIAM Journal on Numerical Analysis* 14 (6): 1006–1021. http://www.jstor.org/stable/2156678.

Anderson, J. 2009. "Spatially and temporally varying adaptive covariance inflation for ensemble filters." *Tellus A: Dynamic Meteorology and Oceanography* 61 (1): 72–83. https://doi.org/10.1111/j.1600-0870.2007.00361.x.

Bastian, P. 2014. "A fully-coupled discontinuous Galerkin method for two-phase flow in porous media with discontinuous capillary pressure." *Computational Geosciences* 18, no. 5 (2014): 779–796. https://doi.org/10.1007/s10596-014-9426-y.

Bastian, P., Blatt, M., Dedner, A., Dreier, N.-A., Engwer, C., Fritze, R., Gräser, C., et al. 2020. "The DUNE framework: Basic concepts and recent developments." *Computers & Mathematics with Applications,* https://doi.org/10.1016/j.camwa.2020.06.007.

Bastian, P., Heimann, F., and Marnach, S. 2010. "Generic implementation of finite element methods in the Distributed and Unified Numerics Environment (DUNE)." *Kybernetika* 46 (2): 294–315. Accessed November 20, 2020. https://dml.cz/dmlcz/140745.

Bauser, H. H. 2018. "Knowledge Fusion in Soil Hydrology." Dissertation, Heidelberg University. https://doi.org/10.11588/heidok.00024713.

Bauser, H. H., Berg, D., Klein, O., and Roth, K. 2018. "Inflation method for ensemble Kalman filter in soil hydrology." *Hydrology and Earth System Sciences* 22 (9): 4921–4934. https://doi.org/10.5194/hess-22-4921-2018.

Bauser, H. H., Jaumann, S., Berg, D., and Roth, K. 2016. "EnKF with closed-eye period – towards a consistent aggregation of information in soil hydrology." *Hydrology and Earth System Sciences* 20:4999–5014. https://doi.org/10.5194/hess-20-4999-2016.

Berg, D. 2018. "Particle Filters for Nonlinear Data Assimilation." Dissertation, Heidelberg University. https://doi.org/10.11588/heidok.00025182.

Blatt, M., Burchardt, A., Dedner, A., Engwer, C., Fahlke, J., Flemisch, B., Gersbacher, C., et al. 2016. "The Distributed and Unified Numerics Environment, Version 2.4." *Archive of Numerical Software* 4 (100): 13–29. https://doi.org/10.11588/ans.2016.100.26526.

Botto, A., Belluco, E., and Camporese, M. 2018. "Multi-source data assimilation for physically based hydrological modeling of an experimental hillslope." *Hydrology and Earth System Sciences* 22 (8): 4251–4266. https://doi.org/10.5194/hess-22-4251-2018.

Burbeau, A., Sagaut, P., and Bruneau, C.-H. 2001. "A Problem-Independent Limiter for High-Order Runge–Kutta Discontinuous Galerkin Methods." *Journal of Computational Physics* 169, no. 1 (2001): 111–150. https://doi.org/10.1006/jcph.2001.6718.

Burgers, G., Jan van Leeuwen, P., and Evensen, G. 1998. "Analysis Scheme in the Ensemble Kalman Filter." *Monthly Weather Review* 126 (6): 1719–1724. https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Carsel, R. F. and Parrish, R. S. 1988. "Developing joint probability distributions of soil water retention characteristics." *Water Resources Research* 24 (5): 755–769. https://doi.org/10.1029/WR024i005p00755.

Chaudhuri, A., Hendricks Franssen, H.-J., and Sekhar, M. 2018. "Iterative filter based estimation of fully 3D heterogeneous fields of permeability and Mualem-van Genuchten parameters." *Advances in Water Resources* 122:340–354. https://doi.org/10.1016/j.advwatres.2018.10.023.

Clément, J.-B., Golay, F., Ersoy, M., and Sous, D. 2020. "Adaptive Discontinuous Galerkin Method for Richards Equation." In *Topical Problems of Fluid Mechanics 2020*. Prague, Czech Republic. https://doi.org/10.14311/TPFM.2020.004.

Cockburn, B., Gopalakrishnan, J., and Lazarov, R. 2009. "Unified Hybridization of Discontinuous Galerkin, Mixed, and Continuous Galerkin Methods for Second Order Elliptic Problems." *SIAM Journal on Numerical Analysis* 47, no. 2 (2009): 1319–1365. https://doi.org/10.1137/070706616.

Crestani, E., Camporese, M., Baú, D., and Salandin, P. 2013. "Ensemble Kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation." *Hydrology and Earth System Sciences* 17 (4): 1517–1531. https://doi.org/https://doi.org/10.5194/hess-17-1517-2013.

Dedner, A., Kane, B., Klöfkorn, R., and Nolte, M. 2018. "Python framework for hp-adaptive discontinuous Galerkin methods for two-phase flow in porous media." *Applied Mathematical Modelling* 67:179–200. https://doi.org/10.1016/j.apm.2018.10.013.

Di Pietro, D. A. and Ern, A. 2012. *Mathematical Aspects of Disconinuous Galerkin Methods.* Edited by G. Allaire and J. Garnier. Mathématiques et Applications 69. Berlin and Heidelberg: Springer. https://doi.org/10.1007/978-3-642-22980-0.

Dryja, M. 2003. "On Discontinuous Galerkin Methods for Elliptic Problems with Discontinuous Coefficients." *Computational Methods in Applied Mathematics* 3 (1): 76–85. https://doi.org/10.2478/cmam-2003-0007.

Erdal, D., Rahman, M. A., and Neuweiler, I. 2015. "The importance of state transformations when using the ensemble Kalman filter for unsaturated flow modeling: Dealing with strong nonlinearities." *Advances in Water Resources,* Data assimilation for improved predictions of integrated terrestrial systems, 86:354–365. https://doi.org/10.1016/j.advwatres.2015.09.008.

Ern, A., Nicaise, S., and Vohralík, M. 2007. "An accurate H(div) flux reconstruction for discontinuous Galerkin approximations of elliptic problems." *Comptes Rendus Mathematique* 345 (12): 709–712. https://doi.org/10.1016/j.crma.2007.10.036.

Ern, A. and Stephansen, A. F. 2008. "A posteriori energy-norm error estimates for advection-diffusion equations approximated by weighted interior penalty methods." *Journal of Computational Mathematics* 26 (4): 488–510. Accessed November 20, 2020. https://www.jstor.org/stable/43693458.

Ern, A., Stephansen, A. F., and Zunino, P. 2009. "A discontinuous Galerkin method with weighted averages for advection–diffusion equations with locally small and anisotropic diffusivity." *IMA Journal of Numerical Analysis* 29:235–256. https://doi.org/10.1093/imanum/drm050.

Ern, A. and Vohralík, M. 2009. "Flux reconstruction and a posteriori error estimation for discontinuous Galerkin methods on general nonmatching grids." *Comptes Rendus Mathematique* 347 (7): 441–444. https://doi.org/10.1016/j.crma.2009.01.017.

Evensen, G. 1994. "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics." *Journal of Geophysical Research: Oceans* 99 (C5): 10143–10162. https://doi.org/10.1029/94JC00572.

Evensen, G. 2003. "The Ensemble Kalman Filter: theoretical formulation and practical implementation." *Ocean Dynamics* 53 (4): 343–367. https://doi.org/10.1007/s10236-003-0036-9.

Farthing, M. W. and Ogden, F. L. 2017. "Numerical Solution of Richards' Equation: A Review of Advances and Challenges." *Soil Science Society of America Journal* 81, no. 6 (2017): 1257–1269. https://doi.org/10.2136/sssaj2017.02.0058.

Ferracina, L. and Spijker, M. 2005. "An extension and analysis of the Shu-Osher representation of Runge-Kutta methods." *Mathematics of Computation* 74 (249): 201–219. https://doi.org/10.1090/S0025-5718-04-01664-3.

Geuzaine, C. and Remacle, J.-F. 2009. "Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities." *International Journal for Numerical Methods in Engineering* 79 (11): 1309–1331. https://doi.org/10.1002/nme.2579.

Gevaert, A. I., Teuling, A. J., Uijlenhoet, R., DeLong, S. B., Huxman, T. E., Pangle, L. A., Breshears, D. D., et al. 2014. "Hillslope-scale experiment demonstrates the role of convergence during two-step saturation." *Hydrology and Earth System Sciences* 18, no. 9 (2014): 3681–3692. https://doi.org/https://doi.org/10.5194/hess-18-3681-2014.

Hamill, T. M., Whitaker, J. S., and Snyder, C. 2001. "Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter." *Monthly Weather Review* 129 (11): 2776–2790. https://doi.org/10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

Hendricks Franssen, H.-J. and Kinzelbach, W. 2008. "Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem." *Water Resources Research* 44 (9). https://doi.org/10.1029/2007WR006505.

Hendrickx, J. and Flury, M. 2001. "Uniform and Preferential Flow Mechanisms in the Vadose Zone." In *Conceptual Models of Flow and Transport in the Fractured Vadose Zone*, 149–187. Washington, D.C.: National Academy Press, 2001. https://doi.org/10.17226/10102.

Houtekamer, P. L. and Mitchell, H. L. 2001. "A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation." *Monthly Weather Review* 129 (1): 123–137. https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Ippisch, O., Vogel, H. .-., and Bastian, P. 2006. "Validity limits for the van Genuchten–Mualem model and implications for parameter estimation and numerical simulation." *Advances in Water Resources* 29 (12): 1780–1789. https://doi.org/10.1016/j.advwatres.2005.12.011.

Ippisch, O. 2016. "Contributions to the large-scale Simulation of Flow and Transport in Heterogeneous Porous Media." Habilitation, Heidelberg University. https://doi.org/10.11588/heidok.00020488.

Jaumann, S. and Roth, K. 2017. "Effect of unrepresented model errors on estimated soil hydraulic material properties." *Hydrology and Earth System Sciences* 21, no. 9 (2017): 4301–4322. https://doi.org/10.5194/hess-21-4301-2017.

Jury, W. A., Or, D., Pachepsky, Y., Vereecken, H., Hopmans, J. W., Ahuja, L. R., Clothier, B. E., et al. 2011. "Kirkham's Legacy and Contemporary Challenges in Soil Physics Research." *Soil Science Society of America Journal* 75 (5): 1589–1601. https://doi.org/https://doi.org/10.2136/sssaj2011.0115.

Klein, O. 2016. *DUNE-Randomfield.* Computer software. Version 2.6. Accessed November 10, 2020. https://gitlab.dune-project.org/oklein/dune-randomfield.

Lawson, W. G. and Hansen, J. A. 2004. "Implications of Stochastic and Deterministic Filters as Ensemble-Based Data Assimilation Methods in Varying Regimes of Error Growth." *Monthly Weather Review* 132, no. 8 (2004): 1966–1981. https://doi.org/10.1175/1520-0493(2004)132<1966:IOSADF>2.0.CO;2.

Leeuwen, P. J. v. 2009. "Particle Filtering in Geophysical Systems." *Monthly Weather Review* 137, no. 12 (2009): 4089–4114. https://doi.org/10.1175/2009MWR2835.1.

Li, H., Farthing, M. W., and Miller, C. T. 2007. "Adaptive local discontinuous Galerkin approximation to Richards' equation." *Advances in Water Resources* 30 (9): 1883–1901. https://doi.org/10.1016/j.advwatres.2007.02.007.

List, F. and Radu, F. A. 2016. "A study on iterative methods for solving Richards' equation." *Computational Geosciences* 20, no. 2 (2016): 341–353. https://doi.org/10.1007/s10596-016-9566-3.

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., et al. 2012. "Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities." *Hydrology and Earth System Sciences* 16 (10): 3863–3887. https://doi.org/10.5194/hess-16-3863-2012.

Liu, Y. and Gupta, H. V. 2007. "Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework." *Water Resources Research* 43 (7). https://doi.org/10.1029/2006WR005756.

Miller, C. T., Dawson, C. N., Farthing, M. W., Hou, T. Y., Huang, J., Kees, C. E., Kelley, C. T., and Langtangen, H. P. 2013. "Numerical simulation of water resources problems: Models, methods, and trends." *Advances in Water Resources* 51:405–437. https://doi.org/10.1016/j.advwatres.2012.05.008.

Miller, E. E. and Miller, R. D. 1956. "Physical Theory for Capillary Flow Phenomena." *Journal of Applied Physics* 27 (4): 324–332. https://doi.org/10.1063/1.1722370.

Mualem, Y. 1976. "A new model for predicting the hydraulic conductivity of unsaturated porous media." *Water Resources Research* 12 (3): 513–522. https://doi.org/10.1029/WR012i003p00513.

Oki, T. and Kanae, S. 2006. "Global Hydrological Cycles and World Water Resources." *Science* 313, no. 5790 (2006): 1068–1072. https://doi.org/10.1126/science.1128845.

Ospina De Los Ríos, S. 2019. "Conservative High Order Methods for the Solute Transport Equation in Unsaturated Porous Media." Master Thesis, Heidelberg University.

Pachepsky, Y. and Hill, R. L. 2017. "Scale and scaling in soils." *Geoderma,* Soil structure and function in a changing world: Characterization and scaling, 287:4–30. https://doi.org/10.1016/j.geoderma.2016.08.017.

Reichle, R. H. 2008. "Data assimilation methods in the Earth sciences." *Advances in Water Resources,* Hydrologic Remote Sensing, 31, no. 11 (2008): 1411–1418. https://doi.org/10.1016/j.advwatres.2008.01.001.

Riedel, L. 2017. "Advancing Data Assimilation of Soil Water Flow with a DG-based Richards Solver." Master Thesis, Heidelberg University.

Roth, K. 1995. "Steady State Flow in an Unsaturated, Two-Dimensional, Macroscopically Homogeneous, Miller-Similar Medium." *Water Resources Research* 31 (9): 2127–2140. https://doi.org/https://doi.org/10.1029/95WR00946.

Roth, K. 2008. "Scaling of water flow through porous media and soils." *European Journal of Soil Science* 59 (1): 125–130. https://doi.org/10.1111/j.1365-2389.2007.00986.x.

Roth, K. 2012. *Soil Physics.* Lecture Notes. Heidelberg, Germany: Institute of Environmental Physics, Heidelberg University. Accessed September 10, 2020. http://ts.iup.uni-heidelberg.de/fileadmin/user_upload/misc/teaching/sp-2.2.pdf.

Schroeder, W., Martin, K., and Lorensen, B. 2006. *The Visualization Toolkit.* 4th ed. Edited by A. Squillacote. Kitware, Inc. Accessed November 20, 2020. https://www.kitware.com/products/books/VTKTextbook.pdf.

Scudeler, C., Paniconi, C., Pasetto, D., and Putti, M. 2017. "Examination of the seepage face boundary condition in subsurface and coupled surface/subsurface hydrological models." *Water Resources Research* 53 (3): 1799–1819. https://doi.org/10.1002/2016WR019277.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J. 2010. "Investigating soil moisture–climate interactions in a changing climate: A review." *Earth-Science Reviews* 99, no. 3 (2010): 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004.

Šimůnek, J., Jarvis, N. J., van Genuchten, M. T., and Gärdenäs, A. 2003. "Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone." *Journal of Hydrology,* Soil Hydrological Properties and Processes and their Variability in Space and Time, 272, no. 1 (2003): 14–35. https://doi.org/10.1016/S0022-1694(02)00252-4.

Sivapalan, M. and Blöschl, G. 2017. "The Growth of Hydrological Understanding: Technologies, Ideas, and Societal Needs Shape the Field." *Water Resources Research* 53 (10): 8137–8146. https://doi.org/https://doi.org/10.1002/2017WR021396.

Söderlind, G. 2006. "Time-step selection algorithms: Adaptivity, control, and signal processing." *Applied Numerical Mathematics,* Selected Papers, The Third International Conference on the Numerical Solutions of Volterra and Delay Equations, 56, no. 3 (2006): 488–502. https://doi.org/10.1016/j.apnum.2005.04.026.

Solin, P. and Kuraz, M. 2011. "Solving the nonstationary Richards equation with adaptive hp-FEM." *Advances in Water Resources,* New Computational Methods and Software Tools, 34 (9): 1062–1081. https://doi.org/10.1016/j.advwatres.2011.04.020.

The HDF Group. 1997–2020. *Hierarchical Data Format, Version 5.* Computer software. Accessed November 19, 2020. http://www.hdfgroup.org/HDF5/.

Van Genuchten, M. T. 1980. "A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils." *Soil Science Society of America Journal* 44:892–898. https://doi.org/10.2136/sssaj1980.03615995004400050002x.

Van Leeuwen, P. J., Cheng, Y., and Reich, S. 2015. "Nonlinear Data Assimilation for high-dimensional systems." In *Nonlinear Data Assimilation.* Frontiers in Applied Dynamical Systems: Reviews and Tutorials. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-18347-3_1.

Van Schaik, N. L. M. B., Schnabel, S., and Jetten, V. G. 2008. "The influence of preferential flow on hillslope hydrology in a semi-arid watershed (in the Spanish Dehesas)." *Hydrological Processes* 22 (18): 3844–3855. https://doi.org/10.1002/hyp.6998.

Vereecken, H., Schnepf, A., Hopmans, J. W., Javaux, M., Or, D., Roose, T., Vanderborght, J., et al. 2016. "Modeling Soil Processes: Review, Key Challenges, and New Perspectives." *Vadose Zone Journal* 15, no. 5 (2016). https://doi.org/10.2136/vzj2015.09.0131.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., et al. 2020. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." *Nature Methods* 17:261–272. https://doi.org/10.1038/s41592-019-0686-2.

Vogel, H.-J. 2019. "Scale Issues in Soil Hydrology." *Vadose Zone Journal* 18 (1). https://doi.org/10.2136/vzj2019.01.0001.

Vogel, H.-J. and Roth, K. 2003. "Moving through scales of flow and transport in soil." *Journal of Hydrology,* Soil Hydrological Properties and Processes and their Variability in Space and Time, 272, no. 1 (2003): 95–106. https://doi.org/10.1016/S0022-1694(02)00257-3.

Vogel, T., van Genuchten, M. T., and Cislerova, M. 2000. "Effect of the shape of the soil hydraulic functions near saturation on variably-saturated flow predictions." *Advances in Water Resources* 24 (2): 133–144. https://doi.org/10.1016/S0309-1708(00)00037-3.

Vrugt, J. A., Stauffer, P. H., Wöhling, T., Robinson, B. A., and Vesselinov, V. V. 2008. "Inverse Modeling of Subsurface Flow and Transport Properties: A Review with New Developments." *Vadose Zone Journal* 7 (2): 843–864. https://doi.org/10.2136/vzj2007.0078.

Wu, C.-C. and Margulis, S. A. 2011. "Feasibility of real-time soil state and flux characterization for wastewater reuse using an embedded sensor network data assimilation approach." *Journal of Hydrology* 399 (3): 313–325. https://doi.org/10.1016/j.jhydrol.2011.01.011.

Zeng, X. and Decker, M. 2009. "Improving the Numerical Solution of Soil Moisture–Based Richards Equation for Land Models with a Deep or Shallow Water Table." *Journal of Hydrometeorology* 10, no. 1 (2009): 308–319. https://doi.org/10.1175/2008JHM1011.1.

Zhang, Q., Shi, L., Holzman, M., Ye, M., Wang, Y., Carmona, F., and Zha, Y. 2019. "A dynamic data-driven method for dealing with model structural error in soil moisture data assimilation." *Advances in Water Resources* 132:103407. https://doi.org/10.1016/j.advwatres.2019.103407.