Numerical simulation of flows in unsaturated porous media by an adaptive discontinuous Galerkin method: application to sandy beaches

Jean-Baptiste Clément

UNIVERSITÉ DE TOULON

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULON

Délivré par
Université de Toulon

Discipline ou spécialité
Mécanique des fluides, Mathématiques appliquées

Présentée et soutenue par

Jean-Baptiste CLÉMENT

Le 11 janvier 2021

Simulation numérique des écoulements en milieu poreux non-saturés par une méthode de Galerkine discontinue adaptative : application aux plages sableuses

---

École doctorale
ED 548 : Mer et Sciences

Unités de recherche
Institut de Mathématiques de Toulon – IMATH (EA 2134)
Institut Méditerranéen d'Océanographie – MIO (UMR 7294 CNRS/UTLN/AMU)

Encadrement de thèse

M. Frédéric GOLAY,	Maître de conférences HDR – Univ. de Toulon,	Directeur de thèse
M. Damien SOUS,	Maître de conférences HDR – Univ. de Toulon/UPPA,	Co-directeur de thèse
M. Mehmet ERSOY,	Maître de conférences HDR – Univ. de Toulon,	Co-encadrant

Jury

M. Philippe HELLUY,	Professeur – Université de Strasbourg,	Président du Jury
M. Vít DOLEJŠÍ,	Professeur – Charles University in Prague,	Rapporteur
M. Philippe ACKERER,	Directeur de recherche – CNRS/Univ. de Strasbourg,	Rapporteur
Mme Béatrice RIVIÈRE,	Professeure – Rice University,	Examinatrice
Mme France FLOC'H,	Maître de conférences – UBO,	Examinatrice
M. Stéphane BONELLI,	Directeur de recherche – INRAE,	Invité

PHD THESIS MANUSCRIPT

Numerical simulation of flows in unsaturated porous media by an adaptive discontinuous Galerkin method: application to sandy beaches

Jean-Baptiste CLÉMENT

11th January 2021

Doctoral advisor: Dr. Frédéric GOLAY

Doctoral co-advisor: Dr. Damien SOUS

Co-supervisor: Dr. Mehmet ERSOY

Institut de Mathématiques de Toulon – IMATH (EA 2134), Université de Toulon, Bâtiment M - Campus La Garde, CS 60584, 83041 Toulon Cedex 9

Institut Méditerranéen d'Océanographie – MIO (UMR 7294 CNRS/IRD), Aix-Marseille Université, Université de Toulon, Bâtiment X - Campus La Garde, CS 60584, 83041 Toulon Cedex 9

Á ma mère, mon père et ma sœur.

Foreword – Avant-propos

Abstract – Résumé

Flows in unsaturated porous media are modelled by the Richards' equation which is a degenerate parabolic nonlinear equation. Its limitations and the challenges raised by its numerical solution are laid out. Getting robust, accurate and cost-effective results is difficult in particular because of moving sharp wetting fronts due to the nonlinear hydraulic properties. Richards' equation is discretized by a discontinuous Galerkin method in space and backward differentiation formulas in time. The resulting numerical scheme is conservative, high-order and very flexible. Thereby, complex boundary conditions are included easily such as seepage condition or dynamic forcing. Moreover, an adaptive strategy is proposed. Adaptive time stepping makes nonlinear convergence robust and a block-based adaptive mesh refinement is used to reach required accuracy cost-effectively. A suitable *a posteriori* error indicator helps the mesh to capture sharp wetting fronts which are also better approximated by a discontinuity introduced in the solution thanks to a weighted discontinuous Galerkin method. The approach is checked through various test-cases and a 2D benchmark. Numerical simulations are compared with laboratory experiments of water table recharge/drainage and a largescale experiment of wetting, following reservoir impoundment of the multi-materials La Verne dam. This demanding case shows the potentiality of the strategy developed in this thesis. Finally, applications are handled to simulate groundwater flows under the swash zone of sandy beaches in comparison with experimental observations.

Les écoulements en milieux poreux non-saturés sont modélisés par l'équation de Richards qui est une équation non-linéaire parabolique dégénérée. Ses limites et les défis que soulèvent sa résolution numérique sont présentés. L'obtention de résultats robustes, précis et efficaces est difficile en particulier à cause des fronts de saturation raides et dynamiques induits par les propriétés hydrauliques non-linéaires. L'équation de Richards est discrétisée par une méthode de Galerkine discontinue en espace et des formules de différentiation rétrograde en temps. Le schéma numérique résultant est conservatif, d'ordre élevé et très flexible. Ainsi, des conditions aux limites complexes sont facilement intégrées comme la condition de suintement ou un forçage dynamique. De plus, une stratégie adaptative est proposée. Un pas de temps adaptatif rend la convergence non-linéaire robuste et un raffinement de maillage adaptatif basée sur des blocs est utilisée pour atteindre la précision requise efficacement. Un indicateur d'erreur *a posteriori* approprié aide le maillage à capturer les fronts de saturation raides qui sont également mieux approximés par une discontinuité introduite dans la solution grâce à une méthode de Galerkine discontinue pondérée. L'approche est validée par divers cas-tests et un benchmark 2D. Les simulations numériques sont comparées à des expériences de laboratoire de recharge/drainage de nappe et une expérience à grande échelle d'humidification, suite à la mise en eau du barrage multi-matériaux de La Verne. Ce cas exigeant montre les potentialités de la stratégie développée dans cette thèse. Enfin, des applications sont menées pour simuler les écoulements souterrains sous la zone de jet de rive de plages sableuses en comparaison avec des observations expérimentales.

Keywords – Mots-clés

Richards' equation, seepage, weighted DG method, AMR, *a posteriori* error estimation, BDF, swash groundwater.

Équation de Richards, suintement, méthode GD pondérée, raffinement de maillage, estimation d'erreur *a posteriori*, BDF, hydrodynamique souterraine du jet de rive.

Acknowledgements – Remerciements

First and foremost, I would like to thank the members of my jury to have accepted to evaluate my work during its defence. Thanks to Philippe Helluy for chairing my jury while everything was virtual because of the COVID-19 pandemic. I would like to extend my thanks to Vít Dolejší and Philippe Ackerer for the time spent reviewing my manuscript as well as the remarks that they gave in their report. In particular, thank you, Mr. Dolejší, to have received me in your office for these two fruitful discussions about DG. Thank you also to Béatrice Rivière, France Floc'h and Stéphane Bonelli for their interest in my work and being part of the jury. I was thrilled by meeting you properly to discuss and not being just names on papers that I studied. A special thank to Cédric Galusinski, Professor and director of IMATH, and Stéphane Bonelli who gave me valuable scientific and organization advice during the annual PhD committees. Finally, thank to Prof. Ing. Jaroslav Fořt to have welcomed me in his laboratory from Czech Technical University in Prague a first time for an unforgettable internship and then for two months during my PhD. This was a great support for me!

J'aimerais ensuite remercier mon directeur de thèse, Frédéric Golay, pour son implication dans ce projet, ses conseils scientifiques et par-dessus tout son humanité. Il a su apporter les réponses appropriées pendant les moments difficiles. Je retiens enfin sa pédagogie – notamment en matière de programmation ! – et sa passion pour la Recherche qui m'ont poussé à entreprendre cette thèse alors que j'étais encore étudiant puis qui m'ont guidé jusqu'à ce moment présent qui, j'espère, ne sera pas le dernier. Avec tout le respect et l'affection : merci de m'avoir fait padawan ! Merci aussi à Sylvie pour ses encouragements.

À Damien Sous, mon co-directeur de thèse, dont chaque réunion s'est révélée instructive grâce à ses connaissances sur la physique du littoral et son questionnement toujours juste. J'ai découvert que les plages et les vagues sont un monde de complexité que je ne soupçonnais pas ! Tes interprétations physiques et ton sens rédactionnel resteront dans ma mémoire. Merci de m'avoir fait confiance et de m'avoir encouragé tout au long de ce projet. J'espère pouvoir mener les applications à leur terme dans un futur proche !

À Mehmet Ersoy, mon encadrant, qui aura toujours été là pour me conseiller et m'épauler dans ce projet. Sa grande rigueur et sa passion pour les mathématiques m'auront orienté tout le long de cette thèse et chaque séance de travail en sa compagnie m'ont permis de construire des connaissances solides. Je le aussi remercie pour son humour, son enthousiasme et ses relectures éclairs qui auront été d'une grande aide. Merci à vous : Fred, Dam et M. !

Je remercie tous les chercheurs et personnels de l'IMATH, de COSMER et SeaTech qui ont rendu ces trois années agréables à vivre. En particulier, j'aimerais remercier les collègues doctorants - l'équipe du fameux Bureau des Doctorants ! Aux anciens à qui je souhaite une très bonne continuation : Minh, Hadjer, Moustoifa (quoi que tu fasses !), Houssam, Yssouf et Mohamed Ali. À ceux qui soutiennent bientôt : Asma, Soufiane, courage, vous tenez le bon bout ! Aux nouveaux, je suis très honoré d'avoir fait votre connaissance et j'espère que vous passerez le flambeau pour les générations à venir : Youssef,

Juliette, Bilal, Clémentin, Ali et Viktoriia. To those who were there briefly: Karol and Pepa. Special děkuji to Pepa who shared his hospitality by welcoming me in Prague for two months. I hope you're doing well and good luck for your PhD! Dédicace aux stagiaires : bonne continuation !

Il y a forcément certaines personnes avec qui les atomes ont été plus crochus ! Merci à Hoàng Anh, doctorant hors-classe, qui est devenu au fil de cette thèse un ami formidable, plein d'énergie mais qui est surtout passé maître dans la négociation, l'humour et la cuisine ! Même si je pense pouvoir deviner ton prochain repas du midi ! La thèse n'a été que plus agréable en ta compagnie. Merci Mai Anh, tu es toujours rayonnante, et merci à ton micro de qualité pour la soutenance ! Obrigado à Matheus pour sa gentillesse dépassant l'entendement et son sens de l'amitié. Les discussions autour d'une bière et les sorties plage ont une saveur particulière. C'est tout naturellement que je dis dankeschön à Dorothea dont la joie de vivre, la culture et son regard sur le monde ont également rendu ces sorties spéciales. Je vous souhaite tous deux de vivre pleinement vos projets peu importe où le vent vous mène ! Merci à Manon pour les discussions toujours intéressantes au bureau et les moments de détente. Ça m'a fait beaucoup de bien en particulier les sorties escalade avec Romain ! Je remercie Ornella, la respo bureau, qui a toujours été à mi-cheval de me faire rire comme lorsqu'elle s'est retrouvée coincée dans le piège-à-Ornella. Un grand merci pour son soutien et en particulier son aide à MT180 : ses talents ont fait merveilles pour numériser et coloriser ma planche ! Merci à Nicolas et Anna (et Sacha !) pour leur accueil et leur sens du jeu pendant le temps des confinements. Gate, Mr-2-Tot, tropfzahn, Raoulde-Cambrai, Ornella : thank you very much for the enjoyable games: 14!

Un grand merci aux amis de SeaTech, en particulier la promo MOCA 2017 ; à jamais les meilleurs ! Leurs joutes verbales sont toujours sources de franches rigolades et je souhaite bon courage à ceux qui finalisent leur thèse. Merci aussi à l'équipe de doctorants de MT180 d'avoir rendu ce moment magique.

Comment ne pas penser à la fine équipe : Thibaut, Célestin et Charline, François, Rémi ! Que de bons moments partagés ensemble qui ont permis de jalonner cette thèse ! Thibaut, je souhaite du fond du cœ]ur que tu mènes ta thèse telle que tu la voies. Quant à tes élèves, Célestin, j'espère qu'ils feront beaucoup de thèses ; tu as une mission ! Julien, nous sommes passés à travers l'épreuve du feu de la prépa, et depuis, ça m'a fait plaisir de te revoir au gré de ton parcours international.¡Cindy, no puedo esperar a que me muestres los secretos de México! Felicitaciones por su matrimonio. Enfin, Lucien et Vincent, ce fut à chaque fois très agréable de vous retrouver pour discuter avec vous !

Je souhaite remercier Marcel et Cathy pendant ces trois années de spéléo et canyon qui ont été essentielles à ma santé mentale ! Votre gentillesse et votre passion sont très inspirantes. Je ne pouvais pas rêver mieux que vous pour me faire découvrir cet univers aussi fabuleux que technique. Merci à toutes les personnes que j'ai croisées pendant ces sorties.

Je mesure la chance que j'ai d'avoir rencontré Maxime à qui je dédie des remerciements particulièrement appuyés. Son soutien sans faille et son humour imparable m'ont été d'une grande aide tout au long de cette thèse, sans compter les souvenirs inoubliables : merci, vraiment ! Onze ans, tu te rends compte ? Il est bien loin le temps des TPE !

Je remercie de tout mon cœur ma famille - oncles, tantes, cousins, cousines, grands-parents - qui a toujours été très encourageants ! Je tiens à exprimer toute la gratitude à mon papa, ma maman et ma sœur sans qui cette thèse n'aurait pu se faire. Les prémisses remontent sûrement à loin car c'est vous qui m'aviez donné goût à cette curiosité toujours émerveillée et cette pensée critique toujours renouvelée qui font ce que je suis ! Un grand merci à ma sœur qui est un exemple de courage et une source inépuisable de drôleries !

JB, le 18 janvier 2021

—

I am partially supported by a grant from Provence-Alpes-Côte d'Azur region, France.

Contents

Nomenclature

Acronyms

Acronym	Definition
3MT	Three Minute Thesis
Rivage	RIchards pour des VAgues par méthodes de Galerkine discontinuEs
AGMG	Iterative solution with AGgregation-based algebraic MultiGrid
AMR	Adaptive Mesh Refinement
AMU	Aix-Marseille Université
BARDEX	BARrier Dynamics EXperiment
BDF	Backward Differentiation Formula
CFL	Courant-Friedrichs-Lewy
CNRS	Centre National de la Recherche Scientifique
CPU	Central Processing Unit
CSR	Compressed Sparse Row
CV	Control volume
DIRK	Diagonally Implicit Runge-Kutta
EA	Équipe d'Accueil
ED	École Doctorale
FD	Finite Difference
FE	Finite Element
Fortran	FORmula TRANslation
FV	Finite Volume
HDR	Habilitation à Diriger des Recherches
IIPG	Incomplete Interior Penalty Galerkin
IMATH	Institut de MATHématiques de Toulon
INRAE	Institut National de recherche pour l'Agriculture, l'alimentation et l'Environnement

IRD	Institut de Recherche pour le Développement
MIO	Mediterranean Institute of Oceanography
MT180	Ma Thèse en 180 secondes
NIPG	Non-symmetric Interior Penalty Galerkin
OBB	Oden-Baumann-Babuška
PDE	Partial Differential Equation
RANS	Reynolds Average Navier-Stokes
REV	Representative Elementary Volume
SIPG	Symmetric Interior Penalty Galerkin
SKS	SKyline matrix Storage
SOLLC	SOLveur Ligne de Ciel
SOLP	SOLveur Plein
SWASH	Simulating WAves till SHore
UBO	Université de Bretagne Occidentale
UMR	Unité Mixte de Recherche
USDA	United States Department of Agriculture
UTLN	Université de TouLoN
VOF	Volume Of Fluid
WDG	Weighted Discontinuous Galerkin

Abbreviations

Abbreviation	**Definition**
1D, 2D, 3D	one-dimension, two-dimension, three-dimension
e.g.	*exempli gratia*, for example
et al.	*et alii*, and others
etc	*et cetera*, and so on
i.e.	*id est*, that is
cst	constant
Dr.	Doctor
Eq.	Equation
Fig.	Figure
Ing.	Engineer
M.	Monsieur

Mme	Madame
PhD	*Philosophiae Doctor*, Doctor of Philosophy
Prof.	Professor
Tab.	Table
Univ.	Université
w.r.t.	with respect to

Dimensionless numbers

Symbol	Description	Definition
Fr	Froude number	$\frac{u}{\sqrt{gL}}$
Ma	Mach number	$u\sqrt{\frac{\rho}{p}}$
Pe	Péclet number	$\frac{vL}{D}$
Re	Reynolds number	$\frac{\rho u L}{\mu}$

Superscripts

Superscript	Definition
$\square^*$	Nondimensionalization scales variable
$\square^n$	Time-discretized, at time t^n

Greek symbols

Symbol	Description	Dimension
α	Multi-index	-
α	Pore-size distribution, air entry pressure inverse (parameter)	-
$\alpha_{s,q}$	Coefficients for BDF methods	-
β	Penalty order	-
β_{c}	Threshold value for coarsening	-
β_{m}	Mass balance	-
β_{r}	Threshold value for refinement	-
β_{T}	Parameter for Tracy's solution	-
χ_k^l	Degrees of freedom	-
$\Delta_{\mathbb{K}}$	Degree of anisotropy	-

$\delta_h^{n+1,m}$	Solution increment in the nonlinear iterative process	L
$\partial\Omega$	Domain boundary	-
$\Delta x, \Delta y$	Horizontal and vertical grid spacing	L
ϵ	Nondimensionalized fluid radius	-
η	Error estimate	-
ε	Small number	-
ϵ_0	Nondimensionalized VER radius	-
$\varepsilon_1, \varepsilon_2$	User-defined tolerances	-
η_c	Compactness	-
η_E^n	Estimation-based error indicator	-
$\eta_{\mathrm{F}E}^n$	Face residual of estimation-based error indicator	-
$\eta_{\mathrm{J}E}^n$	Solution jump of estimation-based error indicator	-
ϵ_m	Error on mass balance	-
ϵ_pr	Machine epsilon	-
$\eta_{\mathrm{R}E}^n$	Element residual of estimation-based error indicator	-
γ_F	Diffusion penalty coefficient for a face	-
Γ_D	Dirichlet boundary	-
Γ_N	Neumann boundary	-
Γ_S	Seepage boundary	-
γ_T	Parameter for Tracy's solution	-
κ	Normal component of hydaulic conductivity for a face	-
κ_m	Minimum of neighbouring normal hydaulic conductivities of a face	-
λ_amp	Amplification factor for time stepping	-
λ_k	Parameter for Tracy's solution	-
λ_M	Largest eigenvalue of the hydraulic conductivity tensor	-
λ_m	Lowest eigenvalue of the hydraulic conductivity tensor	-
λ_red	Reduction factor for time stepping	-
μ	Dynamic viscosity	$\mathrm{M}\cdot\mathrm{L}^{-1}\cdot\mathrm{T}^{-1}$
μ_F	Face measure	L
μ_k	Parameter for Tracy's solution	-
Φ	Porosity	-
$\hat{\varphi}_i$	Shape functions	-

Φ_k^l	Global polynomial basis functions	-
ϕ_k^l	Local polynomial basis functions	-
Ψ	Deviation rate of pressure	-
ψ_e	Air entry pressure	L
ψ	Pressure head	L
ψ_ex	Exact solution for Tracy's benchmark	L
ψ_s	Minimum capillary height	L
ψ_ss	Steady-state solution for Tracy's benchmark	L
ψ_ts	Transient counterpart for Tracy's benchmark	L
ρ	Density	$\mathrm{M{\cdot}L^{-3}}$
ϱ_F^D	Penalty weight for a Dirichlet boundary face	$\mathrm{L{\cdot}T^{-1}}$
ρ_E	Element roundness	L
ϱ_F^I	Penalty weight for an interior face	$\mathrm{L{\cdot}T^{-1}}$
ϱ_F^X	Extrapenalty weight for an interior face	$\mathrm{L{\cdot}T^{-1}}$
$\Sigma(x, y)$	Surface of the ball of center x and radius y	-
σ_F^D	Penalty speed for a Dirichlet boundary face	$\mathrm{L{\cdot}T^{-1}}$
σ_E	Element aspect ratio	-
σ_*	Threshold value for the penalty speed	$\mathrm{L{\cdot}T^{-1}}$
σ_F^I	Penalty speed for an interior face	$\mathrm{L{\cdot}T^{-1}}$
τ	Time step	-
τ_m	Error rate on mass balance	-
θ	Water content	-
ϑ	Parameter for the ϑ-method	-
Θ	Symmetry weight	-
θ_r	Residual water content	-
θ_s	Saturated water content	-
Θ_S	Transformation function preserving the surface	-
Θ_V	Transformation function preserving the volume	-
Υ	Deviation rate of velocity	-
Ω	Domain	-
ω	Weight for WDG	-
ξ	Physical quantity	-

Subscripts

Subscript	Definition
$\square_{\text{air}}$	Air phase
$\square_{\alpha}$	α-phase
$\square_{,\text{B}}$	Boundary faces
$\square_{\text{bottom}}$	At the bottom
$\square_{,\text{D}}$	Dirichlet boundary faces
$\square_{\mathcal{E}}$	Energy norm (intrinsic norm)
$\square_{\mathscr{F}}$	Fluid phase
$\square_{h}$	Space-discretized with space step h
$\square_{,\text{I}}$	Interior boundary faces
$\square_{\text{J}}$	Jump seminorm
$\square_{\text{l}}$	Trace on left element
$\square_{,\text{N}}$	Neumann boundary faces
$\square_{\text{new}}$	New mesh element
$\square_{\text{O}}$	Element origin
$\square_{\text{old}}$	Old mesh element
$\square_{\text{R}}$	Residual seminorm
$\square_{\text{r}}$	Trace on right element
$\square_{,\text{S}}$	Seepage boundary faces
$\square_{\mathscr{S}}$	Solid phase
$\square_{\text{top}}$	At the top
$\square_{\text{water}}$	Water phase

Latin symbols

Symbol	Description	Dimension
A	Empirical shape parameter	L^{B}
a, b	Width and height of rectangular domain	L
$a_{h,n}$	Bilinear form for stiffness terms	$\text{L}^{d+1}\cdot\text{T}^{-1}$
A_{ij}	Stiffness matrix	-
$\mathcal{B}(x, r_0)$	Ball of center x and radius r_0	-
B	Empirical shape parameter	-

C	Capillary capacity	L^{-1}
c	Coefficient expressing the fact that a part of the kinetic energy of the real fluid turns into heat	-
$\mathcal{C}^0$	Set of continuous functions	-
C	Empirical shape parameter	L^{D}
C_E	Gradient-based error indicator	-
c_τ, c'_τ	Coefficients independent of τ	-
c_h, c'_h	Coefficients independent of h	-
$\mathbb{D}$	Diffusivity tensor	$\mathrm{L}^2{\cdot}\mathrm{T}^{-1}$
d	Dimensionr	-
D^α	EPartial derivative	-
∂E	Element boundary	-
D	Empirical shape parameter	-
E	Element	-
$\mathcal{E}$	Mesh (set of elements)	-
e_h, e_h^n	True error	-
F	Face	-
F_E	Mapping from reference to physical element	-
$\mathcal{F}$	Set of faces)	-
$f_{h,0}$	Linear form for initial terms	$\mathrm{L}^{d+1}{\cdot}\mathrm{T}^{-1}$
g	Gravitational acceleration	$\mathrm{L}{\cdot}\mathrm{T}^{-2}$
H	Total head	L
h	Hydraulic head	L
h	Space step (maximum element diameter of the mesh)	L
H^s	Hilbert space $W^{s,2}$	-
h_0	Initial hydraulic head	L
h_D	Dirichlet boundary hydraulic head	L
h_E	Element diameter	L
h_F	Face measure	L
h_r	Head loss	L
I	Identity matrix	-
I^eff	Effectivity index	-
J_D	Dirichlet boundary penalty	$\mathrm{L}^{d+1}{\cdot}\mathrm{T}^{-1}$

J_E	Jacobian matrix for the mapping	-
J_{I}	Interior penalty	$\mathrm{L}^{d+1}\cdot\mathrm{T}^{-1}$
J_{X}	Extrapenalty	$\mathrm{L}^{d+1}\cdot\mathrm{T}^{-1}$
$\mathbb{K}$	Hydraulic conductivity tensor	$\mathrm{L}\cdot\mathrm{T}^{-1}$
$\mathbb{k}$	Permeability tensor	L^2
k	Permeability	L^2
K_{r}	Relative hydraulic conductivity	-
k_{r}	Relative permeability	-
$\mathbb{K}_{\mathrm{s}}$	Intrinsic or saturated hydraulic conductivity tensor	$\mathrm{L}\cdot\mathrm{T}^{-1}$
$\mathbb{k}_{\mathrm{s}}$	Intrinsic or saturated permeability tensor	L^2
K_{s}	Saturated hydraulic conductivity	$\mathrm{L}\cdot\mathrm{T}^{-1}$
ℓ	Microscopic characteristic length	L
L	Macroscopic characteristic length	L
l	Pore connectivity	-
$l_{h,n}$	Linear form for load terms	$\mathrm{L}^{d+1}\cdot\mathrm{T}^{-1}$
L_i	Load vector	-
L^p	Lebesgue space	-
m	Mask function	-
m	Tortuosity, pore-size distribution (parameter)	-
m_{add}	Total additional mass	L^{d+1}
m_{art}	Additional artificial mass	L^{d+1}
m_{flux}	Total net mass flux	L^{d+1}
$m_{h,n}$	Bilinear form for mass terms	$\mathrm{L}^{d+1}\cdot\mathrm{T}^{-1}$
M_{it}	Maximum threshold value for time stepping	-
m_{it}	Minimum threshold value for time stepping	-
M_{ij}	Mass matrix	-
$\mathbb{N}$	Set of natural integers	-
$\boldsymbol{n}$	Normal unit derivative	-
N	Number of timer intervals	-
n	Pore-size distribution (parameter)	-
N_{dof}	Number of global degrees of freedom	-
N_{dof}^E	Number of local degrees of freedom	-

N_{G}	Number of Gauss quadrature nodes	-
N_{it}	Number of nonlinear iterations	-
N_{nod}	Number of vertices of an element	-
$\mathbb{P}_p$	Space of polynomial functions of total degree less than or equal to p	-
P	Characteristic pressure	$\mathrm{M}\cdot\mathrm{L}^{-1}\cdot\mathrm{T}^{-2}$
p	Pressure	$\mathrm{M}\cdot\mathrm{L}^{-1}\cdot\mathrm{T}^{-2}$
p	Space approximation order	-
p_0	Atmospheric pressure	$\mathrm{M}\cdot\mathrm{L}^{-1}\cdot\mathrm{T}^{-2}$
P_{c}	Capillary pressure	$\mathrm{M}\cdot\mathrm{L}^{-1}\cdot\mathrm{T}^{-2}$
p_E	Polynomial order for an element	-
p_F	Polynomial order for a face	-
$\mathbb{Q}_p$	Space of polynomial functions of degree less than or equal to p	-
Q	Source/sink term	T^{-1}
q	Flow, Darcy velocity	$\mathrm{L}\cdot\mathrm{T}^{-1}$
q	Time approximation order	$\mathrm{L}\cdot\mathrm{T}^{-1}$
q_{N}	Neumann boundary flux	$\mathrm{L}\cdot\mathrm{T}^{-1}$
$\mathbb{R}$	Set of real numbers	-
r	Radius for the fluid ball preserving volume	L
r_0	Mesoscopic characteristic length	L
$r_{h,n}$	Bilinear form for the residual	$\mathrm{L}^{d+1}\cdot\mathrm{T}^{-1}$
S	Saturation	-
s	Fractional index of Sobolev space	-
s	Radius for the fluid ball preserving surface	L
S_{a}	Aqueous phase saturation	-
S_{e}	Effective saturation	-
S_{e}^*	Normalized effective saturation	-
S_p	Supbspace of a broken Sobolev space	-
S_{s}	Specific storage	L^{-1}
T	Final time, time interval	T
t	Time, time point	T
U	Characteristic fluid velocity	$\mathrm{L}\cdot\mathrm{T}^{-2}$
u	Fluid velocity	$\mathrm{L}\cdot\mathrm{T}^{-2}$

V	Volume	L^3
v	Flow speed at a point on a streamline	$\mathrm{L}\cdot\mathrm{T}^{-1}$
V_{poral}	Poral space	L^3
V_{total}	Total space of the porous medium	L^3
W_{it}	Maximum bound for time stepping	-
w_i	Gauss quadrature weights	-
$W^{s,p}$	Sobolev spave	-
x	Global space variable	L
$\boldsymbol{x}, \boldsymbol{y}$	Space vector	L
x	Lobal space variable	L
z	Point elevation, total space variable	L

Other symbols

Symbol	Definition
$\mathbb{1}(\bullet)$	Indicator function
$\{\!\{\bullet\}\!\}$	Arithmetic average over a face
$\{\!\{\bullet\}\!\}_{\omega}$	Weighted average over a face
$\{\!\{\bullet\}\!\}_{\overline{\omega}}$	Conjugate weighted average over a face
$\widetilde{\bullet}$	Spatial deviation
$\overline{\bullet}$	Intrinsic average
$[\![\bullet]\!]$	Jump over a face
$< \bullet >$	Phase average
$\hat{\bullet}$	Reference element

Introduction

Predicting variably-saturated flows in porous media is a major issue for many fields in science and engineering. For example, such flows arise in soil physics, hydrogeology, environment, agriculture or oil industry for problems like subsurface contaminant transport, petroleum reservoir, water resources, surface ponding, *etc*. Experimental studies are a fundamental tool to describe such processes and have a good understanding. However, experimental studies are not always possible, nor feasible. Mathematical modelling and numerical simulation provide another possibility to explore and gain new insights into such physical systems.

Context and motivations

Understanding the dynamics of groundwater flow in coastal areas is of major interest in the fields of coastal engineering and sustainable development. In the present work, a particular attention will be paid on nearshore groundwater applications. A better understanding of the groundwater circulation is important for analysing and predicting a range of physical and bio-geochemical processes in coastal zones, such as sediment transport (bed stability), the diffusion of dissolved materials such as pollutants or nutrients, or mixing between continental (fresh) and marine (salt) water. Each of these issues requires special attention in the context of climate change and erosion [1, 2] (sea level rise, submergence events, salinisation of arable land) and increasing anthropogenic pressure (coastal urbanisation). Photos of Fig. 1 show examples of such issues.

(a) Waikiki beach, Hawaii (b) Daytona beach, Florida

Figure 1 – Examples of coastal urbanization and threat of sea level rise for beaches with high socio-economic services.

Recent laboratory studies [3, 4] and field studies [5, 6] have provided good experimental knowledge

of the physical processes governing groundwater flow dynamics in sandy beaches exposed to wave action. Recently obtained data during the ROUSTY2014 campaign will allow further analysis of these underground dynamics on a larger scale [7]. Despite their primary importance, experimental measurements remains very costly and are limited in time and space resolution and extent. A major scientific challenge is to develop robust and efficient numerical models to better anticipate, via the realisation of predictive scenarios, the evolution of the coastal zone. Few models have been presented so far and their scope of application remains limited, see *e.g.* [8, 9]. Further numerical developments are needed to address the key issues of beach groundwater induced by the waves.

Many porous media flow models have been proposed and developed to simulate lots of different flows and transport processes at various scales. Multiphase flow problems present a more complex behaviour than single phase flow problem since phase dynamics influences each other. In this context, multiphase flow model may lead to solving difficulties or have a level of details too complex for the range of aimed applications. Simplified model considering reasonable hypotheses are very attractive alternative.

In the present work, flows in variably-saturated porous media are described by Richards' equation. What makes Richards' equation very appealing is that it models the porous medium as a whole part including both saturated and unsaturated zones. Despite neglecting air-phase, Richards' equations accounts for actions both from gravity and capillarity and so can model various and complex processes at different scales. Richards' equation is a nonlinear parabolic equation which can degenerate into an elliptic equation under complete saturation condition. The history of Richards' equation begins with Darcy's law formulated experimentally by Darcy in 1856 [10] for saturated porous media. Then, this result was extended to multiphase flows by Buckingham in 1907 [11] to give the Darcy-Buckingham law, which serves as cornerstone for the derivation of Richards' equation. The latter was first established by Richardson in 1922 [12] but was attributed solely to Richards who published independently the equation in 1931 [13]. First attempts to solve Richards' equation numerically go back to the late 1960s with Rubin [14] or Cooley [15] for example. Then, from 1980s, Richards' equation was studied extensively, both from theoretical and numerical point of views. Results of existence, regularity and uniqueness were found [16] and solution was assessed through simulations involving different numerical methods. One of the points of interest was nonlinear solver [17, 18]. Since 2000s, works focus mainly on Richards' equation stability and development of coupled and general simulations [19]. In recent years, research consists in Richards' equation with hysteresis and advanced spatial/temporal discretizations [20, 21] with adaptive approximations in size and/or in order.

Richards' equation has been extensively used for numerical simulations by the hydrogeology community, see *e.g.* [19, 22]. Despite this research effort, many numerical challenges remain for solving Richards' equation with some simulations being still unreliable and/or expensive [23, 24]. This numerical complexity prevented a more general use of Richards-based model for a number of applications. Indeed, the solution of Richards' equation involves sharp wetting fronts which evolve both in space and time and are difficult to resolve. Besides, the simulation must treat simultaneously unsaturated/saturated regions of parabolic/elliptic natures, strong heterogeneity and anisotropy of the porous medium and possibly fast-changing boundary conditions. These aspects can also lead to steep gradients acting like discontinuities. Considering the set of nonlinear functions for hydraulic properties, it is seldom possible to foresee the behaviour of Richards' equation and getting numerical solutions can be hard to achieve. Especially, numerical schemes often fail to converge, requiring a careful numerical treatment, or need fine discretization, making computation costly.

Objectives and scope of the thesis

The objective of this thesis is to develop tools leading to a simulation of Richards' equation which is:

- accurate, in the sense that the numerical solution should be physically acceptable and converge to an *exact* solution when time steps and/or mesh resolution is refined;
- robust, in the sense that the method should provide a solution whatever the physical and numerical parameters for the problem are which means convergence reliability;
- efficient, in the sense that computation should be cost-effective for time and memory resources;
- suitable for beach groundwater induced by the waves.

The main targeted application of the present numerical research work is the simulation of groundwater flows within sandy beaches. Luijendijk *et al.* [1] found that more than one-third of the world's ice-free coastline are sandy. Satellite data analysis over a period of 33 years (1984-2016) indicates that a quarter of sandy beaches are eroding at rates exceeding 0.5 m/yr. Sandy shorelines have high socio-economic benefits related to tourism and ecosystem services. They are a serious concern because of climate change since sea level rise put sandy beaches under threat of increasing erosion as Vousdoukas *et al.* asserted [2]. But Cooper *et al.* [25] qualified the disappearance of sandy beaches which can retreat inland. Nevertheless, sandy beaches remain subject to huge change and may disappear if urbanization blocks their migration. Besides, from a modelling point of view, sandy beaches are usually dissipative beach with weak slope where groundwater flow is known to be slow with a weak retroaction on the inducing waves. Then, a weak coupling of surface flow on groundwater flow can be considered to target applicative simulations. These reasons make the thesis prioritize the study of groundwater for sandy beaches.

For this purpose, a discontinuous Galerkin (DG) method is chosen to solve Richards' equation. DG methods are based on a variational formulation in an element-wise fashion, sharing advantages both with finite elements and finite volumes methods. In particular, they are locally conservative which is crucial in fluid dynamics [26]. Moreover, the nature of DG formulation enables to work on non-conforming mesh and to change locally the degree of polynomial approximation. This is an important benefit since Adaptive Mesh Refinement (AMR), the so-called h-adaptation, and high-order accuracy, the so-called p-adaptation, become possible [27], promoting a growing use of DG methods for transport phenomena in porous media such as two-phase flow problems [28, 29].

Besides, Richards' equation is known to be a stiff differential equation with difficult convergence [17, 18, 30] so the time discretization and the nonlinear solver have to be addressed carefully.

Keeping in mind that the numerical method should provide accurate and robust results, but efficient in terms of computational time, this thesis is an opportunity to contribute to update high-order adaptive DG methods for simulation of Richards' equation. Then, the following achievements are pursued:

- The Richards' equation is discretized by primal DG methods which lead to mass conservative, high-order and AMR to get accurate solution. Penalization weakly enforces boundary conditions in the DG formulation for robustness. Backward differentiation formulas are used for time discretization because they are known to be high-order and stable for stiff equation;
- A fixed-point method together with adaptive time stepping make nonlinear iterations more reliable;

- An *a posteriori* error indicator is formulated for Richards' equation approximated by DG methods. This estimate is broken down into different contributions which guide mesh adaptation;
- The algorithm employs block-based AMR to save computational time.

Contributions and outline

This thesis proposes a novel strategy to investigate adaptive simulations in time and space for Richards' equation. AMR will be used to capture moving wetting fronts thanks to *a posteriori* estimation. *h*-Adaptation is employed for two-phase flow in porous media [28] but also for Richards' equation [20, 31]. Making the most of DG methods flexibility, AMR is combined with a weighted discontinuous Galerkin (WDG) framework which allows discontinuity in the solution according to the nonlinear diffusivity. Following earlier works, this approach was formulated by Ern, Di Pietro and other collaborators [32, 33] and by Proft and Rivière [34, 35]. Application of such strategy for Richards' equation, combining AMR, *a posteriori* estimate and WDG framework, is the main novelty of the study.

A dedicated code called *Rivage* has been designed, developed and benchmarked during this thesis. *Rivage* will serve in the laboratory for further study of Richards' equation but will also extended to other problems. Additional cases based on laboratory and largescale experiments are under construction and analysis.

The thesis is organized as follows.

Chapter I introduces Richards' equation in a comprehensive manner to get the model problem. Porous media are described to introduce basis notions and phenomenology of porous media flows. Upscaling methods are discussed as modelling tools to obtain fundamental relations for flow through porous media. In particular, they can serve to better understand the use of Richards' equation and are interesting working perspectives. The derivation of Richards' equation is given to draw the limitations of the model. Sections are devoted to hydraulic properties and the seepage boundary condition. They approximate specific processes. Besides, they hold nonlinearities responsible to numerical difficulties for Richards' equation which are discussed.

In Chapter II, Richards' equation is discretized. After introducing DG methods abilities, the space discretization is detailed. The use of weak penalization to enforce boundary conditions is highlighted. Time discretization is performed though implicit methods called backward differentiation formula. The mass balance property of DG methods is exhibited. The nonlinear discrete problem is then solved by an iterative process based on Newton-Raphson method or fixed-point method. An heuristic adaptive time stepping is used to make nonlinear convergence more robust. At the end of this chapter, some implementations choices are provided given that DG methods offer several possibilities.

Chapter III is dedicated to the adaptive strategy proposed to improve solving of Richards' equation. The block-based AMR is presented and justified for Richards' equation with a use of general meshes. Then, a gradient-based error indicator and estimation-based error indicator are introduced and motivated to drive mesh adaptation. They are designed to capture sharp fronts. Some variations for the numerical scheme are also proposed. In particular, jumps and flux are handled by the WDG method to better approximate sharp fronts with a discontinuity.

Several numerical tests are performed in Chapter IV. Numerical experiments allow to get some insights about DG methods. The model is validated against test-cases for 1D vertical infiltration involving different hydraulic properties and boundary conditions. Some specific behaviours of Richards'

equation are noted and performance of mesh adaptation and weighted framework to capture wetting fronts are demonstrated. Tracy's benchmark provides an exact solution which is used to assess convergence and error indicators effectivity. The capabilities of the code are tested on laboratory experiments involving seepage, recharge and drainage for water table. Finally, the modelling strategy is tested on a large-scale challenging case with experimental dataset: the wetting of a multi-materials dam.

In Chapter V, some applications are made for groundwater in sandy beaches. After presenting notions about beach system involving forcing waves and groundwater circulation, comparisons are carried for three cases. An idealized beach is considered to see if the model is able to render main characteristics of beach groundwater observed in experiments from the domain.

At last, a section is devoted to some concluding remarks and perspectives for future works.

Chapter I

Modelling of flows in unsaturated porous media

"La science est faite d'erreurs, mais d'erreurs qu'il est bon de commettre, car elles mènent peu à peu à la vérité."

— Jules Verne, *Voyage au centre de la Terre*, 1864

Notions of hydrogeology and an historical perspective of study of flow in porous media are provided in this chapter. In addition, some insights are given to better understand how it is possible to derive macroscopic flow relations for porous media from fundamental equations at pore scale. This overview enables to set useful concepts in order to describe the governing equation used in the next parts of the thesis to model unsaturated porous media flow: the Richards' equation. Constitutive laws and specific boundary conditions are also introduced to complete the model. Richards' equation is a well-known model, broadly employed for simulations of infiltration in many fields with a long background. However, this situation may be misleading for new arrivals in the domain because underlying hypotheses for its use are often ignored. Besides, solving Richards' equation may be very difficult largely due to nonlinearities of hydraulic properties. This is particularly the case when one wants to simulate non-standard phenomenon or reach accuracy, robustness and cost-effectiveness. So, solution

of Richards' equation is discussed to highlight flow mechanisms that rise numerical challenges as well as computational developments commonly used to deal with them. A qualitative analysis is presented to attempt to sort Richards' equation behaviour.

I.1 Description of porous media

Most of the definitions and notions related here can be found in textbooks about flow in porous media, for instance [22, 36, 37].

A **porous medium** is a material constituted of voids filled by fluids, either liquids or gas. The solid portion of material is called the **skeleton** or matrix while the void portions are called **pores**. Lots of natural and artificial structures can be considered as porous media. Besides, porous media are the place of many processes leading to many concepts for different fields of studies such as mechanics, geosciences, biology, material science or engineering. In particular, fluid flow through porous media is a subject of interest since it explains lots of observations made for natural structures or industrial systems. For example, knowing the way in which fluid flows through a porous medium is useful to describe water infiltration into a soil column or fluid moving in an absorbing sponge. A porous medium is characterised by many variables. **Porosity** is an essential property. It is a number between 0 and 1 which is the measure of pore volume over the total volume. Figure I.1 illustrates porous media with different porosity and pore structures. Indeed, other properties are implied to describe fully porous flow phenomena like mechanical dispersion: tortuosity (path length), viscosity (pore friction) and velocity (pore section). Then, porous media can have the same porosity but they exhibit different pore connectivity and types of porosity.

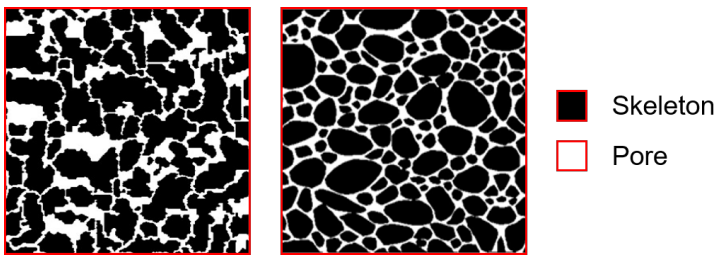


Figure I.1 – Schematic representation of a part from two different porous media.

Porous media present multi-scale structure. Indeed, pores may have various sizes and shapes resulting in different flow regimes. Fluid flow are mainly driven by two kinds of forces. Firstly, if pores are small enough, then liquid is displaced thanks to intermolecular forces. It is made possible by the combination of cohesive forces within the liquid (surface tension) and adhesive forces between the liquid and the surrounding solid matrix. This is called **capillarity**. Secondly, flow in large pores are subject to gravitational forces. Capillarity effects can be sufficiently strong to elevate fluid such as in plants to mo ve water up or paper towels to absorb liquid.

A **phase** is a chemically homogeneous portion of the porous medium that is separated from other such portions by a definite physical boundary. Solid phase represents the skeleton while pores are filled by fluid phases. Fluid phases are immiscible if they maintain a boundary between them. Fluid flow in porous media can be either one-phase flow or multiphase flow. In this latter case, two-phase flow are particularly studied. To simplify description, a two-phase flow with water and air is considered from now on. Single- and multi-phase flow are outside the scope of the thesis. Most of the time, they can

be figured out from the following description but some additional phenomena may rise like osmotic forces or chemical reactions between different phases. In a two-phase system, there is a wetting (water) and a non-wetting (air) phase. **Wettability** definition is linked to the contact angle between fluid phases and the solid skeleton and plays a role in capillary forces.

Saturation represents the quantity of water contained in a porous medium. It is the volume fraction of water compared to the pore volume and ranges from 0 to 1. In practice, the range is smaller because of trapped air bubbles or liquid film around grains. In this case, residual saturation and maximum saturation defines a new normalized quantity called effective saturation. The porous medium is saturated if water fills fully the pore space. Otherwise, the porous medium is said to be unsaturated. Then, two different zones are distinguished: the *saturated zone* (or phreatic zone) and the *unsaturated zone* (or vadose zone). The *water table* is the surface delimiting these two zones. The *capillary fringe* is generally considered as a part of the unsaturated zone. It is the layer in which water seeps up from water table due to capillary forces. Capillary fringe exhibits a nearly-saturated state which may extend from few centimetres to several meters. Capillary fringe size depends strongly on local porosity. Figure I.2 illustrate the different zones of saturation as discussed above. Capillary fringe is showcased with an artesian well in which water cannot rise above water table with capillary forces.

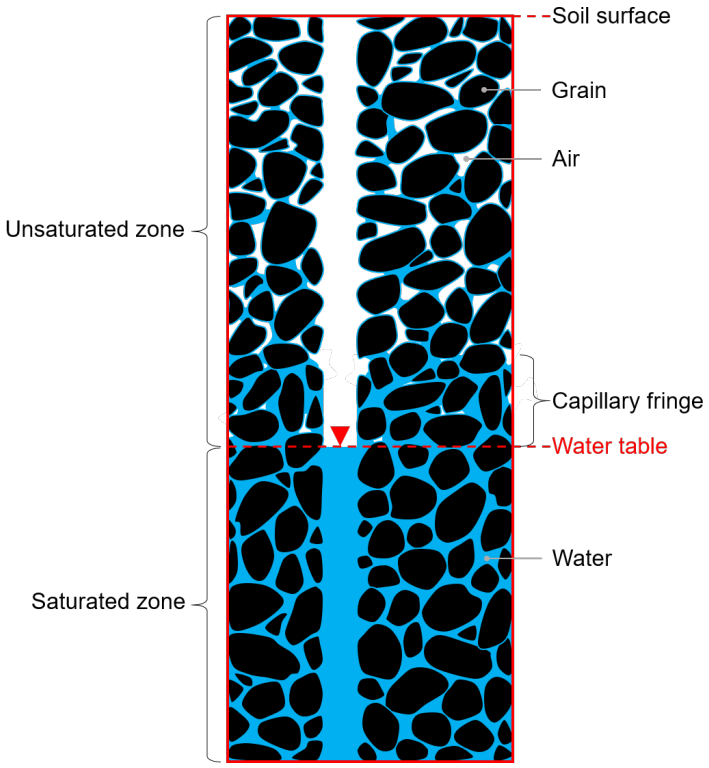


Figure I.2 – The different zones of saturation in a porous medium.

A large part of study of flow in porous media concerns water movements in soils. Geologist have a specific set of definitions to describe **granulometry**. The classification showed in Fig. I.3 is a simplified version from the classification adopted by the United States Department of Agriculture (USDA). One should notice that type names refers to grain size but it does not mean that the particle is composed of this mineral even though it is often the case. Grain size can be used to guess how water infiltrates soil because it is linked to pore size which determines capillarity. However, other parameters contribute to water dynamics like pore shape, pore continuity or pore distribution.

Grain diameter range	Type name	Common name
> 256 mm	Boulder	Stone
64 mm – 256 mm	Cobble	
4 mm – 64 mm	Pebble	Gravel
2 mm – 4 mm	Granule	
0.5 mm – 2 mm	Coarse sand	Sand
0.25 mm – 0.5 mm	Medium sand	
1/16 mm – 0.25 mm	Fine sand	
1/256 mm – 1/16 mm	Silt	Mud
< 1/256 mm	Clay	

Figure I.3 – Scale of grain size.

Soil texture focuses on the particles that are less than two millimeters in diameter which include sand, silt and clay, see Fig. I.4. Proportion of these three particles in the soil defines a class of texture. The classification of soil texture can serve to assess water movements qualitatively within soils because texture is related to chemical and physical properties. It is a useful tool for agriculture, in particular to know soil's capacity to retain water. USDA developed a system of twelve soil textures which is largely used worldwide. Figure I.5 presents the soil texture triangle from USDA.

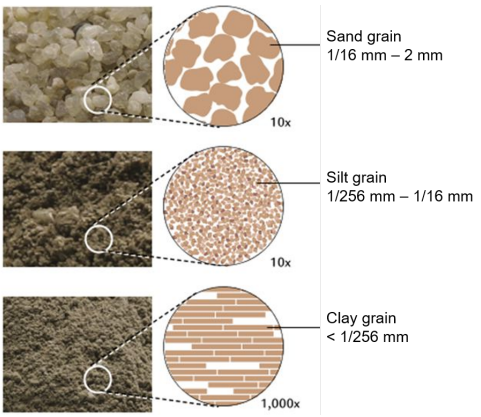


Figure I.4 – Photos with a schematic close-up of a sample of sand-, silt-, clay-sized particles.

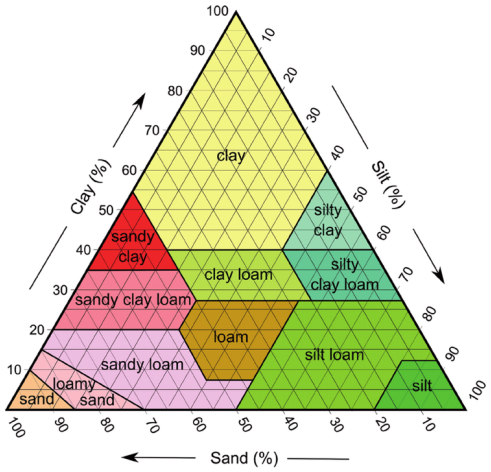


Figure I.5 – Soil texture classification diagram from USDA. Textures are determined by the proportions of sand-, silt-, and clay-sized grains.

Water movement in porous media, and particularly in soils, have different **hydrodynamic regimes** according to capillary and gravitational forces. This regime are linked to different types of water, see Fig. I.6. When air is present, retained water is attached to the grains and its thickness depends on molecular forces. It is made of hygroscopic water, which is in direct contact with the grain, immobile and can only be removed by desiccation or vacuum, and pellicular water, which adheres to hygroscopic water and moves around grains but can be extracted only with centrifugation. Retained water cannot be affected by gravity and does not transmit hydrostatic pressure. Capillary water is subject to tension forces at the interface between air and water. It can be partially extracted by gravity and transmit pressure. Gravitational water is able to be drained out by gravity and its flow moves under gravity and pressure fields. Capillary water and gravitational water are referred as free water and they define flow dynamics in unsaturated porous media. These regimes are sketched in Fig. I.7. Adsorbed regime corresponds to the retained water. At this stage, water can evaporate to make vapour in air phase. In pendular regime, the flow has no continuous paths so is considered as macroscopically immobile [22]. The flow of liquid water is possible from the funicular stage. If the amount of water in the system is sufficient, pores can be filled with water which gives the capillary and saturated regime. Capillary regime contains occluded air bubbles and air phase losses its continuity. Saturated regime happens when remaining air is dissolved into water. From now on, evaporation and dissolution are neglected so we can focus on the funicular, capillary and saturated regimes.

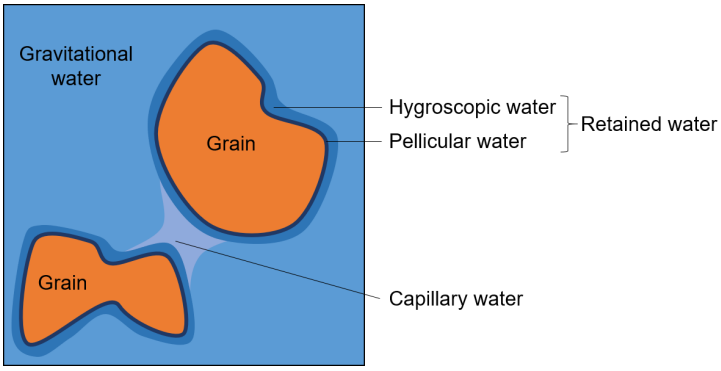


Figure I.6 – Basic concepts for different types of water in soils

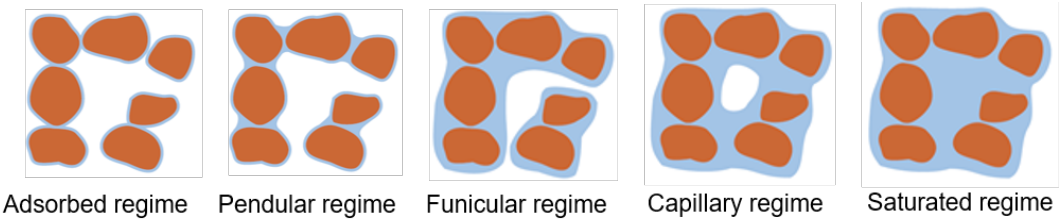


Figure I.7 – Saturation regimes in porous medium showing spatial configurations of air and water. Adapted from Szymkiewicz [22].

The state of water phase is commonly described by the concept of **energy potential**. The total energy is divided into gravitational energy, pressure energy and velocity energy. This is related to the Bernoulli's equation (*i.e.* integrated momentum of the Euler equations), valid at any arbitrary point along a streamline S:

$$H = z + \frac{p}{\rho g} + \frac{cv^2}{2g} + \frac{1}{g}\int_S \partial_t v \, \mathrm{d}s + h_\mathrm{r}, \tag{I.1}$$

where H [L] denotes the total head, z [L] the point elevation with respect to an arbitrary datum, p [M·L^{-1}·T^{-2}] the pressure, ρ [M·L^{-3}] the water density, v [L·T^{-1}] the flow speed at a point on a streamline, g [L·T^{-2}] the acceleration due to gravity, c the coefficient expressing the fact that a part of the kinetic energy of the real fluid turns into heat and h_r [L] the head loss due to the irreversible transformation of mechanical energy into thermal energy through the mechanism of frictional resistance. According to this equation, groundwater flow can be divided in different head components showed in Fig. I.8. Usually, Bernoulli's principle is only applicable for isentropic flows that is to say for both adiabatic and reversible processes. Then, $c = 1$ and $h_\mathrm{r} = 0$. When porous media is considered, pore flow is very slow so velocity head and inertial head are often neglected.

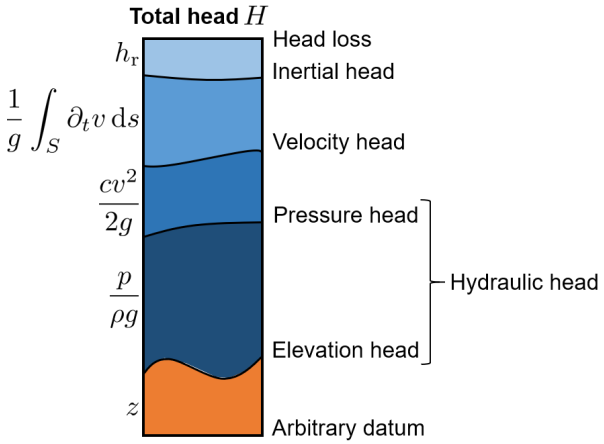


Figure I.8 – Components of the head for water.

I.2 A brief derivation of subsurface flow relation

In this thesis, Richards' equation will be used, see Section I.3. However, Richards' equation derivation lies ultimately on Darcy's law which is the basis relation of flow in porous media. Then, it appears that assumptions to derive Darcy's law are important to know if one wants to employ Richards' equation.

I.2.1 Upscaling approaches

Porous media show heterogeneous structures at microscopic scale inducing spatial variability of their hydraulic properties and characteristics such as porosity, permeability or retention. These microstructures may have an influence at larger scale. However, simulating directly such features would require very fine grids resulting in excessive computational cost. Moreover, from an experimental viewpoint, the microscale properties are difficult to reach or not known with much detail. This makes not purposeful to represent the microscale field if one wants to simulate fields much larger than pore size. Instead, a macroscale representation might be assumed to reproduce the behaviour of the system. Upscaling approaches try to develop such model by the mean of averaged values from a volume element representative of microscale pattern.

Upscaling approches lie on the notion of hierarchy of scales ad spatial variability. Scales are usually separated into a microscopic (local) scale and a macroscopic (large) scale. More scales can be involved, for example, a mesoscopic scale for the volume element or for fractured porous media and even a megascopic scale for aquifers composed of several layers. Spatial variability is fundamental because it often defines the type of upscaling approaches. They can be sorted into three main ideas:

- homogenization methods with, for example, Hornung [38], Allaire [39, 40] or Auriault [41, 42];
- volume average methods with many works from Whitaker [43–46], Quintard [47–51], Gray [52], Neuman [53] or Marle [54, 55];
- stochastic methods with authors like Dagan [56], Zhang [57] or Matheron [58, 59].

Upscaling literature is extensive so a complete review is beyond the scope of the thesis. Consulted studies and general works are listed above. For further insight about authors and contributions of upscaling approaches, Davit *et al.* [60] made a state-of-the-art where a brief history is available and comparison of concepts is carried out like depicted in Fig. I.9. This tremendous amount of contributions leads to many theoretical framework and confusion. Despite evidence of links between these approaches, there is no answer for lots of questions and little effort to clarify/unify these theories as stated by Davit *et al.* [60]. Indeed, upscaling methods are not straightforward since they handle some concepts which may be complex and/or restrictive.

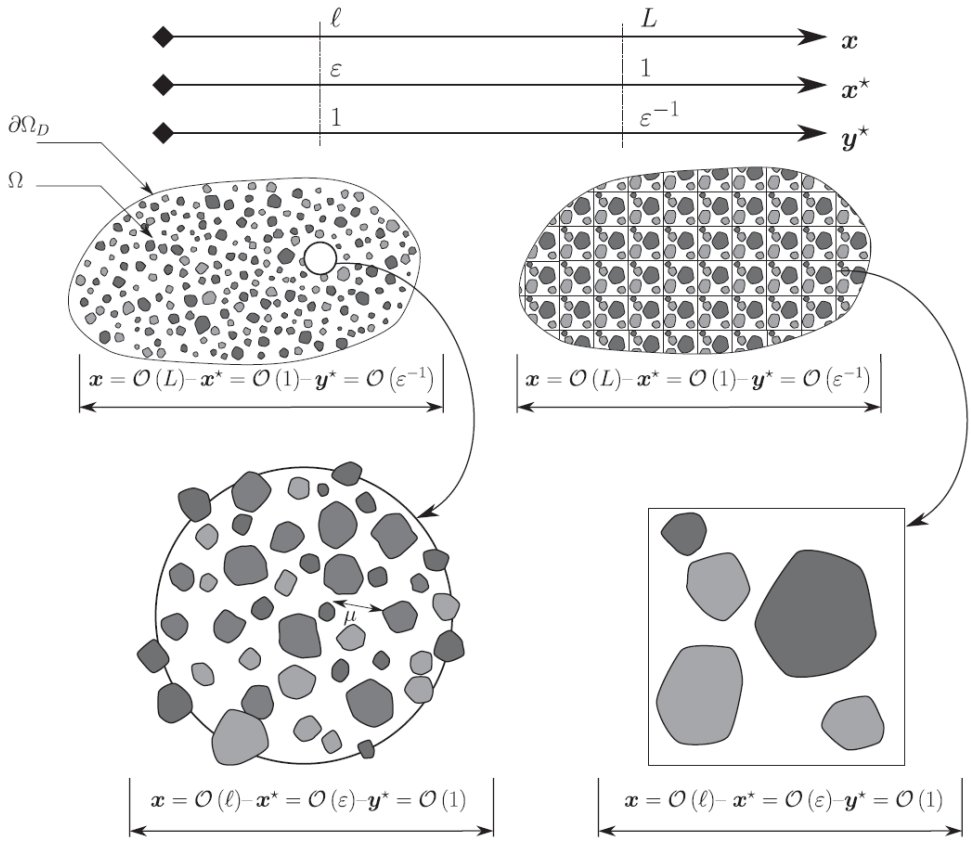


Figure I.9 – Diagram of coordinate systems and scale hierarchy for non-periodic media typical of volume averaging (left) and periodic media typical of homogenization (right). μ is pore-scale characteristic length, ℓ is the size of averaging volume, L is a macroscale characteristic lengths, $\boldsymbol{x}$ is spatial variable coordinate system, $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ are nondimensionalized coordinate systems with L and ℓ respectively. Taken from Davit *et al.* [60].

Homogenization is a very mathematical approach to derive Darcy's law. It is based on asymptotic developments where two scales come into play. Some mechanisms make velocity converge to Darcy's velocity. The derivation is rigorous, well mastered with strong foundation. However, making a physi-

cal interpretation is quite difficult and underlying hypotheses are based on periodic or quasi-periodic structure for porous medium.

Volume average method is a more intuitive approach for Darcy's law derivation. It uses control volume (CV) in which physical quantities are averaged. These averaged quantities are reorganized and some of them are disregarded according to assumptions. The average velocity is obtained and corresponds to the Darcy's velocity. The physical meaning of the process is straightforward. This spontaneous interpretation makes it easy to apply for a wire range of problems. Nevertheless, some assumptions used to neglect terms lack strict justifications and the definition of control volume is not properly established. Indeed, the control volume needs to have the correct size to make averaging meaningful. In this case, control volume is called a representative elementary volume.

Stochastic methods are designed with a statistical and probabilistic point of view. They are less used than the two aforementioned methods but are more fundamental. They work at micro-scale and analyse porous medium properties in a way similar to the ergodic hypothesis.

Due to the variability of spacelengths and timelengths scales of phenomena in porous media, the lack of accurate experimental measures and the wide range of problems, macroscopic description of the system by the mean of upscaling is determining. The boundary effects or the locality of processes define the macroscopic problem which is then used for numerical simulations. A typical example concerns Darcy's law which is a fundamental subsurface flow relation in hydrogeology. Formally, Darcy's law links, on one hand, the flow q [$\mathrm{L}\cdot\mathrm{T}^{-1}$] of a fluid with viscosity μ [$\mathrm{M}\cdot\mathrm{L}^{-1}\cdot\mathrm{T}^{-1}$] trough a porous medium of permeability k [L^2], with, on the other hand, the pressure gradient ∇p [$\mathrm{M}\cdot\mathrm{L}^{-2}\cdot\mathrm{T}^{-2}$]. This relation is explicit and linear:

$$q = -\frac{k}{\mu}\nabla p. \tag{I.2}$$

However, Darcy's law is valid to describe the so-called Darcy flux only if the flow is laminar [36]. Indeed, Darcy's law is the restriction from a much broader relation linking flow velocity with pressure gradient as illustrates in Fig. I.10.

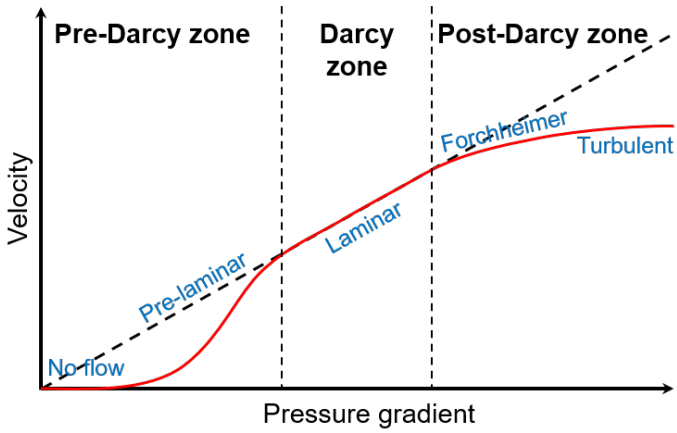


Figure I.10 – Range of validity of Darcy's law. Darcy zone delimits the linear approximation of the flow-pressure gradient relation. Outside, nonlinear deviation are observed.

Darcy's law was originally determined in the 19th century upon experimental results on single-phase flow through a porous medium [10]. Darcy's validity domain is restricted from fine sediments to coarse sand for which pores are small enough compared to the solid skeleton to slow the flow down

to the laminar regime. Usually, typical grain diameter is lower than 1 mm. Otherwise, other laws may be considered, often under the form of a correction. For instance, the Darcy-Forchheimer law adds a quadratic inertial term to Darcy's law to account for nonlinear behaviour when flow is not sufficiently slow to stay laminar. Another correction concerns boundary effects at the interface between a porous medium and free flow when largescale is considered, or, between a grain and a channel pore when microscale is considered. It accounts for viscosity effects within this transition layer, see Fig. I.11 where the Beavers-Joseph-Saffman boundary condition is also shown [61, 62].

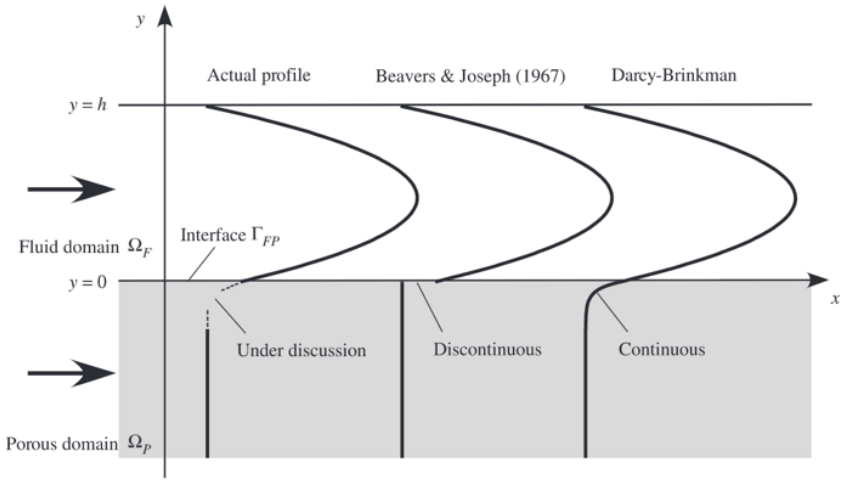


Figure I.11 – Comparison of different interface models for porous media flow. Taken from Fujisawa and Murakami [63].

Subsequent theoretical derivations were undertaken to get Darcy's law or relative forms. Darcy's law can be viewed as a macro-scale representation about what happens at pore-scale where Navier-Stokes or Stokes equations are considered. So, upscaling techniques are particularly suited and can give underlying assumptions about the flow and porous medium natures necessary to derive Darcy's law.

I.2.2 Volume averaging method

Darcy's law is at the basis of the work of this thesis because the same remarks can be formulated when the relation is extended to a two-phase system, *i.e.* using the Darcy-Buckingham law. Indeed, the latter was accepted on experimental and deductive basis. So, as starting point, other derivations for Darcy's law may be proposed for specific problems or highlight "hidden" neglected terms. This makes the upscaling procedure both important and useful to assess the validity of a domain or find new equations. For example, it can be interesting to see effects related to anisotropy, tortuosity, anisotropy, inertia, non-Newtonian flow, *etc*. Upscaling for Darcy-Buckingham law would be a next step since upscaling is more complicated when multiphase flow system is in play [64].

For this thesis, the volume averaging method was chosen to investigate Darcy's law derivation and draw clear hypotheses about its range of validity. Indeed, lots of works are available and its intuitive interpretation may serve as a guide. Many works were read to attempt this goal but those from Whitaker and Quintard were particularly used [43–51]. The process of volume averaging consists in finding a macroscale description of the porous medium from an equation applied at microscale, see Fig. I.12. This transformation is made by the mean of an operator whose application range is called representative elementary volume.

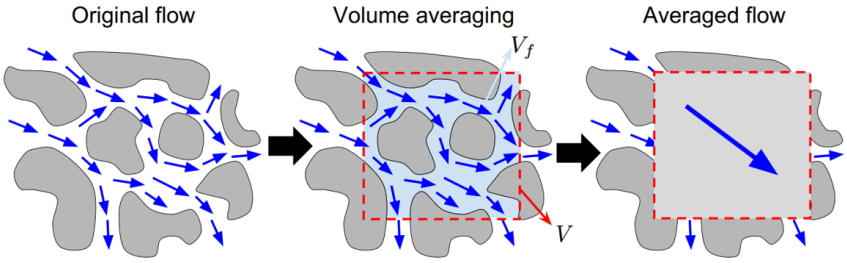


Figure I.12 – Volume averaging procedure to get an equivalent homogeneous macroscale medium from the heterogeneous micro-structure. The red square stands for the representative elementary volume. Adapted from Losada *et al.* [65].

Volume averaging method consists in averaging equations of porous medium over the **Representative Elementary Volume (REV)**. In other words, it is a filtering of small space scales. The general approach aims to average equations to obtain terms which are estimated against each other. Then, given hypotheses at disposal, some of the terms are disregarded. The closure of the problem allows to model non-local terms and so, the rising average velocity is assimilated to the Darcy velocity.

In the following sections, fluid phase is denoted by $\mathscr{F}$ and the solid phase by $\mathscr{S}$. Volume averaging method assumes that there are three space scales satisfying the ordered relation:

$$\ell \leq r_0 \ll L, \tag{I.3}$$

where L corresponds to the macroscopic characteristic length of the porous medium, r_0 is the mesoscopic characteristic length and ℓ is the microscopic characteristic length of the porous structures inside the medium. Let $x \in \Omega \subset \mathbb{R}^3$ be the global space variable sweeping the whole part of the porous medium and y the local space variable for the REV of center x. The space variable z is such that $z = x + y$. The notations, coordinates and scales in play are illustrated in Fig. I.13.

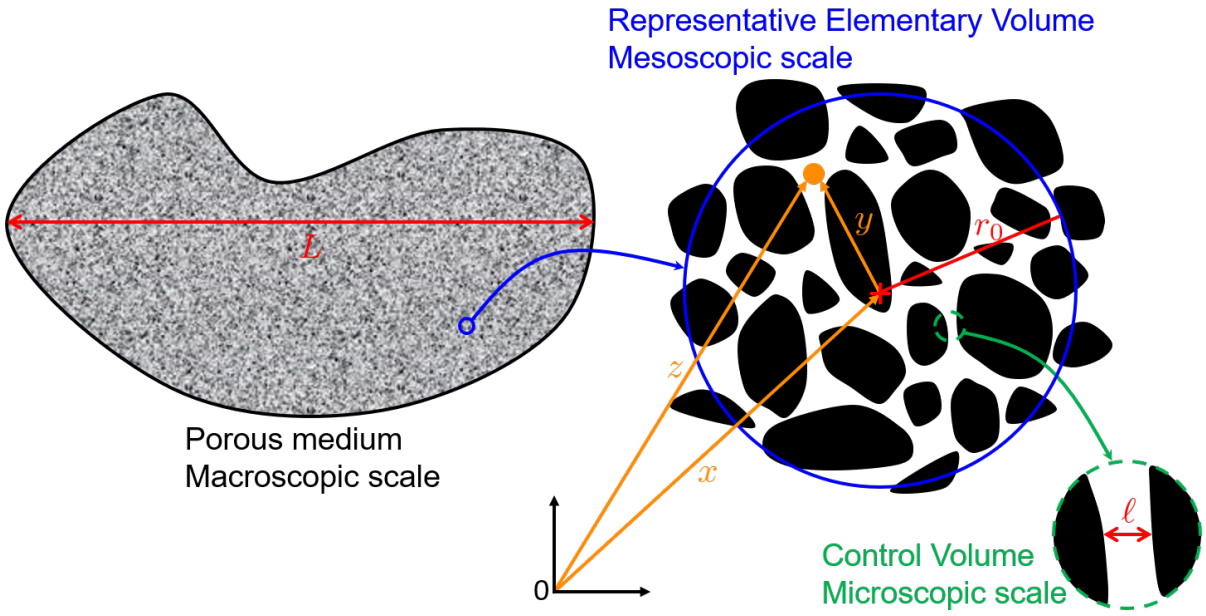


Figure I.13 – Illustration of the coordinate system, notations and hierarchy of scale used for the volume averaging method used in this thesis.

REV size is crucial because it is determining in the averaging process of the quantities but also because it is defining the nature of the porous medium. Yet, it is difficult to give a precise value in practice. Indeed, the REV radius is usually set as the limit of the control volume radius for which the considered variable does not hold fluctuations induced by heterogeneities from small scales, see Fig. I.14. This definition is intuitive but there are several issues like the choice of the defining variable, the control volume shape or the intrinsic nature of the porous medium. Indeed, the porous medium may give rise to different scales where these conditions are met or even no appropriate scale at all. Despite this, the convenience of such a definition makes it useful.

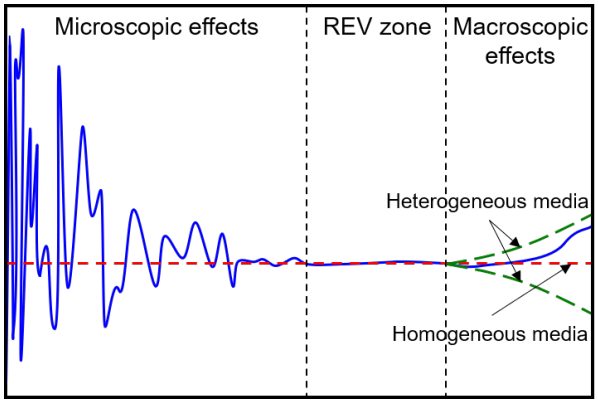


Figure I.14 – Concept for a definition of the Representative Elementary Volume (REV).

The VER can be assimilated into a ball of centre x and radius r_0 so its definition is:

$$\forall x \in \Omega, V = \mathcal{B}(x, r_0) = V_{\mathscr{F}}(x) + V_{\mathscr{S}}(x) = \text{Cst}, \tag{I.4}$$

where we can deduce that $|V| \propto r_0^3$. Similarly, inside a VER, we assume that the solid and fluid phases can be transformed thanks to a suitable function to correspond to balls whose definition is the following:

$$\forall x \in \Omega, \exists \Theta_V \mid \begin{cases} V_{\mathscr{F}}(x) \underset{\Theta_V}{\longmapsto} \mathcal{B}((x, r_{\mathscr{F}}(x)), \\ |V_{\mathscr{F}}(x)| = |\mathcal{B}((x, r_{\mathscr{F}}(x))|, \end{cases} \tag{I.5}$$

$$\forall x \in \Omega, \exists \Theta_S \mid \begin{cases} V_{\mathscr{F}}(x) \underset{\Theta_S}{\longmapsto} \mathcal{B}(x, s_{\mathscr{F}}(x)), \\ |\partial V_{\mathscr{F}}(x)| = |\mathcal{B}(x, s_{\mathscr{F}}(x))|, \end{cases} \tag{I.6}$$

where Θ_V and Θ_S are the transformation functions which preserve respectively the volume and the surface.

For this work, the considered variable to define micro-structures will be the porosity which is a classic choice:

$$\Phi(x) = \frac{V_{\text{poral}}(x)}{V_{\text{total}}}, \tag{I.7}$$

where V_{poral} is the poral space and V_{total} is the total space of the porous medium.

Specific configurations can be distinguished thanks to porosity and characteristic lengths (I is the

identity matrix):

$$\text{If } \ell \sim r_0 \ll L, \text{ then the porous medium is said to be ordered.} \tag{I.8}$$
$$\text{If } \ell \ll r_0 \ll L, \text{ then the porous medium is said to be disordered.} \tag{I.9}$$
$$\text{If } \nabla\Phi \ll I, \text{ then the porous medium is said to be homogeneous.} \tag{I.10}$$
$$\text{If } \nabla\Phi \gg I, \text{ then the porous medium is said to be heterogeneous.} \tag{I.11}$$

When the porous medium is ordered, it means that it follows a periodic or quasi-periodic structure. Homogeneization is often preferred in this case. The more general hypothesis (I.9) is essential for the volume averaging method: it is called the scales separation hypothesis.

Let $\mathbb{1}_{\mathscr{F}}$ be the function defining the fluid phase:

$$\mathbb{1}_{\mathscr{F}}(x+y) = \begin{cases} 1 & \text{if } |y| \leq r_{\mathscr{F}}(x), \\ 0 & \text{otherwise.} \end{cases} \tag{I.12}$$

Since the work focus only the fluid phase, the subscript $\mathscr{F}$ will be omitted when it is not necessary.

We define two filtering operators. The first one is the phase average and for a quantity ξ, it writes:

$$\langle\xi\rangle(x) = \frac{1}{{r_0}^3}\int_{\mathcal{B}(0,r_0)} \mathbb{1}_{\mathscr{F}}(y)\xi(x+y)\,\mathrm{d}y \tag{I.13}$$

The second is the intrinsic average and writes:

$$\overline{\xi}(x) = \frac{1}{r^3(x)}\int_{\mathcal{B}(0,r_0)} \mathbb{1}_{\mathscr{F}}(y)\xi(x+y)\,\mathrm{d}y \tag{I.14}$$

We have the following relationship:

$$\langle\xi\rangle(x) = \Phi(x)\overline{\xi}(x), \tag{I.15}$$

where $\Phi(x) = \dfrac{r^3(x)}{{r_0}^3}$. Besides, it follows:

$$\langle\mathbb{1}_{\mathscr{F}}\rangle(x) = \frac{1}{{r_0}^3}\int_{\mathcal{B}(0,r_0)} \mathbb{1}_{\mathscr{F}}(y)\mathbb{1}_{\mathscr{F}}(x+y)\,\mathrm{d}y = \frac{1}{{r_0}^3}\int_{\mathcal{B}(0,r_0)} \mathbb{1}_{\mathscr{F}}(y)\,\mathrm{d}y = \Phi(x) \tag{I.16}$$

We know that:

$$\begin{aligned} & 0 \leq r(x) \leq r_0 \text{ and } 0 < \ell \ll r_0 \ll L, \\ \iff & 0 \leq \frac{r(x)}{r_0} \leq 1 \text{ and } \frac{\ell}{L} = \epsilon_\ell \ll \frac{r_0}{L} \ll 1, \\ \implies & 0 \leq \epsilon(x) \leq \epsilon_0 \ll 1, \end{aligned} \tag{I.17}$$

where $\epsilon(x) = \dfrac{r(x)}{L}$ and $\epsilon_0 = \dfrac{r_0}{L}$.

Spatial averaging make use of an important result called the spatial averaging theorem [66] which means for a quantity ξ:

$$\langle\nabla\xi\rangle = \nabla\overline{\xi} + \frac{1}{{r_0}^3}\int_{\Sigma(0,s(x))} \mathbb{1}_{\mathscr{F}}(y)\xi(x+y)\,\mathrm{d}y \tag{I.18}$$

where $\Sigma(0, s(x))$ is the surface of the ball of center 0 and radius $s(x)$ corresponding to $\mathcal{B}(0, r(x))$.

To begin the averaging process, we have to choose a local problem to describe the flow at microscale. Seen that flow is very often sufficiently slow in a porous medium, Navier-Stokes equation is simplified into the Stokes equation with a no slip condition:

$$-\mu\Delta u + \nabla p - \rho g = 0, \tag{I.19}$$
$$u \cdot \boldsymbol{n} = 0. \tag{I.20}$$

The phase average is applied to the Stokes equation in Eq. (I.19) and the spatial averaging theorem is used when it is possible. To facilitate visualization of calculation steps, viscous, pressure and gravity terms are treated separately. So it reads:

Pressure terms:

$$\begin{aligned}\langle\nabla p\rangle(x) &= \nabla\langle p\rangle + \frac{1}{{r_0}^3}\int_{\Sigma(0,s(x))} \mathbb{1}_{\mathscr{F}}(y)p(x+y)\boldsymbol{n}\,\mathrm{d}\sigma \\ &= \Phi(x)\nabla\overline{p} + \overline{p}\nabla\Phi(x) + \frac{1}{{r_0}^3}\int_{\Sigma(0,s(x))} \mathbb{1}_{\mathscr{F}}(y)p(x+y)\boldsymbol{n}\,\mathrm{d}\sigma,\end{aligned} \tag{I.21}$$

Viscous terms:

$$\begin{aligned}\langle -\mu\Delta u\rangle(x) &= -\mu\langle\nabla\cdot\nabla u\rangle(x) \\ &\quad - \mu\nabla\cdot\langle\nabla u\rangle - \frac{\mu}{{r_0}^3}\int_{\Sigma(0,s(x))} \mathbb{1}_{\mathscr{F}}(y)(\nabla u\cdot\boldsymbol{n})(x+y)\,\mathrm{d}\sigma \\ &= -\mu\Delta\langle u\rangle + \underbrace{\nabla\cdot\left[-\frac{\mu}{{r_0}^3}\int_{\Sigma(0,s(x))} u(x+y)\boldsymbol{n}\,\mathrm{d}\sigma\right]}_{=0 \text{ with Eq. (I.20)}} \\ &\quad - \frac{\mu}{{r_0}^3}\int_{\Sigma(0,s(x))} \mathbb{1}_{\mathscr{F}}(y)(\nabla u\cdot\boldsymbol{n})(x+y)\,\mathrm{d}\sigma \\ &= -\mu(\Phi(x)\Delta\overline{u} + \overline{u}\Delta\Phi(x) + 2\nabla\Phi(x)\nabla\overline{u}) \\ &\quad - \frac{\mu}{{r_0}^3}\int_{\Sigma(0,s(x))} \mathbb{1}_{\mathscr{F}}(y)(\nabla u\cdot\boldsymbol{n})(x+y)\,\mathrm{d}\sigma,\end{aligned} \tag{I.22}$$

Gravity terms:

$$\langle -\rho g\rangle(x) = -\rho g\langle 1\rangle(x) = -\rho g\Phi(x). \tag{I.23}$$

In the next step, variables will be decomposed with perturbation analysis to study the influence of small scales compared to a reference state. The most used decomposition is the so-called Gray's decomposition. For a variable ξ, it writes:

$$\xi = \overline{\xi} + \widetilde{\xi}. \tag{I.24}$$

Gray's decomposition consists to separate a variable into an intrinsic average component, accounting for slow variations, according to L, and a spatial deviation component, driven by fast variations, according to ℓ. The principle of the process is illustrated in Fig. I.15.

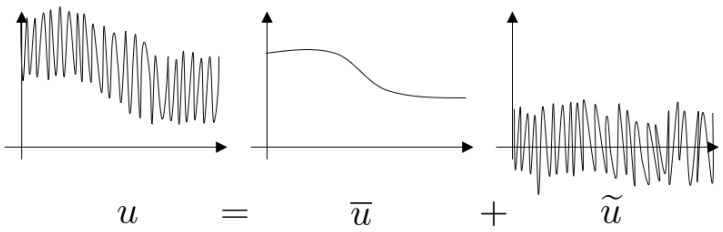


Figure I.15 – Principle of Gray's decomposition between average and deviation fields.

Gray's decomposition is applied on Eq. (I.21) and Eq. (I.22) for u and p which gives:

Pressure terms:

$$\begin{aligned}\langle\nabla p\rangle(x) =&\Phi(x)\nabla\overline{p}+\overline{p}\nabla\Phi(x)\\ &+\frac{1}{{r_0}^3}\int_{\Sigma(0,s(x))}\mathbb{1}_{\mathscr{F}}(y)\overline{p}(x+y)\boldsymbol{n}\,\mathrm{d}\sigma+\frac{1}{{r_0}^3}\int_{\Sigma(0,s(x))}\mathbb{1}_{\mathscr{F}}(y)\widetilde{p}(x+y)\boldsymbol{n}\,\mathrm{d}\sigma\end{aligned} \tag{I.25}$$

Viscous terms:

$$\begin{aligned}\langle-\mu\Delta u\rangle(x) =&-\mu(\Phi(x)\Delta\overline{u}+\overline{u}\Delta\Phi(x)+2\nabla\Phi(x)\nabla\overline{u})\\ &-\frac{\mu}{{r_0}^3}\int_{\Sigma(0,s(x))}\mathbb{1}_{\mathscr{F}}(y)(\nabla\overline{u}\cdot\boldsymbol{n})(x+y)\,\mathrm{d}\sigma\\ &-\frac{\mu}{{r_0}^3}\int_{\Sigma(0,s(x))}\mathbb{1}_{\mathscr{F}}(y)(\nabla\widetilde{u}\cdot\boldsymbol{n})(x+y)\,\mathrm{d}\sigma\end{aligned} \tag{I.26}$$

The following step is a major stage of averaging process. It includes several techniques whose goal is to reject some terms and get a simplified problem by making assumptions for timescales/lengthscales and by approximating deviations. Finally, the last step will obtain a macroscopic equation from closure of the simplified problem. In order to investigate the method of volume averaging and make clear hypotheses about upscaling, a part of the thesis was dedicated to derive a different way to simplify the problem by giving clear assumptions. However, the averaging process could not be concluded during this thesis.

The main choice in this work was to process a nondimensionalization by the mean of several characteristic physical quantities, in particular two different characteristic space lengths. The dimensionless numbers of Reynolds (Re), Mach (Ma) and Froude (Fr) are introduced:

$$\mathrm{Re}=\frac{\rho L\overline{U}}{\mu}, \tag{I.27}$$

$$\mathrm{Ma}=\overline{U}\sqrt{\frac{\rho}{\overline{P}}}, \tag{I.28}$$

$$\mathrm{Fr}=\frac{\overline{U}}{\sqrt{Lg}}. \tag{I.29}$$

Nondimensionalization scales variables with the following intrinsic units to give:

$$\overline{u}^*=\frac{\overline{u}}{\overline{U}},\qquad \overline{p}^*=\frac{\overline{p}}{\overline{P}},\qquad \overline{x}^*=\frac{x}{L}, \tag{I.30}$$

$$\widetilde{u}^*=\frac{\widetilde{u}}{\widetilde{U}},\qquad \widetilde{p}^*=\frac{\widetilde{p}}{\widetilde{P}},\qquad \overline{y}^*=\frac{y}{r_0}. \tag{I.31}$$

However, the star in superscript as well as dependence on x will be omitted in the following equations to preserve readability. Moreover, several quantities which may correspond respectively to deviation rate of velocity, deviation rate of pressure and compactness, are defined:

$$\Upsilon = \frac{\widetilde{U}}{\overline{U}}, \qquad \Psi = \frac{\widetilde{P}}{\overline{P}}, \qquad \eta_{\mathrm{c}}(x) = \frac{s^2(x)}{{r_0}^2}. \tag{I.32}$$

So, we proceed to nondimensionalization and multiply by $\dfrac{L}{\rho\overline{U}^2}$ which gives:

Pressure terms:

$$\begin{aligned}
\langle\nabla p\rangle(x) =& \Phi(x)\frac{\overline{P}}{L}\nabla\overline{p} + \frac{\overline{P}}{L}\overline{p}\nabla\Phi(x) + \frac{\overline{P}s^2(x)}{{r_0}^3}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)\overline{p}(x+s(x)y)\boldsymbol{n}\,\mathrm{d}\sigma \\
&+ \frac{\widetilde{p}s^2(x)}{{r_0}^3}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)\widetilde{p}(x+s(x)y)\boldsymbol{n}\,\mathrm{d}\sigma \\
\langle\nabla p\rangle(x) =& \Phi(x)\frac{\overline{P}}{\rho\overline{U}^2}\nabla\overline{p} + \frac{\overline{P}}{\rho\overline{U}^2}\overline{p}\nabla\Phi(x) + \frac{\overline{P}Ls^2(x)}{\rho{r_0}^3\overline{U}^2}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)\overline{p}(x+s(x)y)\boldsymbol{n}\,\mathrm{d}\sigma \\
&+ \frac{\widetilde{p}Ls^2(x)}{\rho{r_0}^3\overline{U}^2}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)\widetilde{p}(x+s(x)y)\boldsymbol{n}\,\mathrm{d}\sigma \\
\langle\nabla p\rangle(x) =& \frac{1}{\mathrm{Ma}^2}\Phi(x)\nabla\overline{p} + \frac{1}{\mathrm{Ma}^2}\overline{p}\nabla\Phi(x) + \frac{\eta_{\mathrm{c}}(x)}{\epsilon_0\mathrm{Ma}^2}\underbrace{\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)\overline{p}(x+s(x)y)\boldsymbol{n}\,\mathrm{d}\sigma}_{\text{non-local term}} \\
&+\frac{\Psi\eta_{\mathrm{c}}(x)}{\epsilon_0\mathrm{Ma}^2}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)\widetilde{p}(x+s(x)y)\boldsymbol{n}\,\mathrm{d}\sigma
\end{aligned} \tag{I.33}$$

Viscous terms:

$$\begin{aligned}
\langle -\mu\Delta u\rangle(x) =& -\frac{\mu\overline{U}}{L^2}(\Phi(x)\Delta\overline{u} + \overline{u}\Delta\Phi(x) + 2\nabla\Phi(x)\nabla\overline{u}) \\
&- \frac{\mu s^2(x)\overline{U}}{L{r_0}^3}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)(\nabla\overline{u}\cdot\boldsymbol{n})(x+s(x)y)\,\mathrm{d}\sigma \\
&- \frac{\mu s^2(x)\widetilde{U}}{{r_0}^4}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)(\nabla\widetilde{u}\cdot\boldsymbol{n})(x+s(x)y)\,\mathrm{d}\sigma \\
\langle -\mu\Delta u\rangle(x) =& -\frac{\mu}{\rho L\overline{U}}(\Phi(x)\Delta\overline{u} + \overline{u}\Delta\Phi(x) + 2\nabla\Phi(x)\nabla\overline{u}) \\
&- \frac{\mu s^2(x)}{\rho{r_0}^3\overline{U}}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)(\nabla\overline{u}\cdot\boldsymbol{n})(x+s(x)y)\,\mathrm{d}\sigma \\
&- \frac{\mu s^2(x)L\widetilde{U}}{\rho{r_0}^4\overline{U}^2}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)(\nabla\widetilde{u}\cdot\boldsymbol{n})(x+s(x)y)\,\mathrm{d}\sigma \\
\langle -\mu\Delta u\rangle(x) =& -\frac{1}{\mathrm{Re}}\underbrace{(\Phi(x)\Delta\overline{u} + \overline{u}\Delta\Phi(x) + 2\nabla\Phi(x)\nabla\overline{u})}_{\text{Brinkman's correction}} \\
&- \frac{\eta_{\mathrm{c}}(x)}{\epsilon_0\mathrm{Re}}\underbrace{\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)(\nabla\overline{u}\cdot\boldsymbol{n})(x+s(x)y)\,\mathrm{d}\sigma}_{\text{non-local term}} \\
&-\frac{\Upsilon\eta_{\mathrm{c}}(x)}{{\epsilon_0}^2\mathrm{Re}}\int_{\Sigma(0,1)} \mathbb{1}_{\mathscr{F}}(s(x)y)(\nabla\widetilde{u}\cdot\boldsymbol{n})(x+s(x)y)\,\mathrm{d}\sigma
\end{aligned} \tag{I.34}$$

Gravity terms:

$$\begin{aligned} \langle -\rho g \rangle(x) &= -\frac{Lg}{\overline{U}^2}\Phi(x) \\ \langle -\rho g \rangle(x) &= -\frac{1}{\mathrm{Fr}^2}\Phi(x) \end{aligned} \tag{I.35}$$

The averaged equation corresponds to the boxed terms in yellow. One should notice that there are terms referred to the Brinkman's correction taking into account some boundary effects with inertia. Besides, non-local terms are present: dimensional analysis should go on with these terms by performing asymptotic developments related to the quantity $x + s(x)y$. We would obtain terms related to moments of different orders which can be estimated. Then, the terms highlighted in yellow should be kept in order to get an averaged equation defining Darcy's velocity. However, as said earlier, the different estimated weights are difficult to compare making these terms difficult to neglect. Finally, the deviation terms would be replaced by a closure problem to derive the Darcy's law.

This nondimensionalization is not standard compared to the one found in literature. Indeed, the core idea is to estimate the quantities in regard of several characteristic lengths following the concept of scale separation. However, the process was not able to complete successfully during the thesis. This prolongation constitutes a serious perspective of research whose goal would be to derive subsurface flow relations adapted according to dimensionless quantities defining different hydrodynamic regimes.

For now, this derivation stays an open problem but it clarified the understanding about underlying assumptions for Darcy's law which is the fundamental relation on which Richards' equation derivation lies. Limitations involve porous medium structure (scales separation hypotheses, existence of a VER), creeping flow (Stokes equation), no sleep condition with grains (boundary condition), phase equilibrium (surface terms) and porous medium/flow features (Brinkman's correction, Gray's decomposition, non-local terms, *etc*).

I.3 Governing equation

Richards' equation is a classic nonlinear parabolic equation used to describe flow occurring in unsaturated/saturated zones of an aquifer. This section is devoted to a synthetic recall of the Richards' equation derivation, which is rather sparse in the literature [22, 67]. In a porous medium, mass conservation principle stands for two-phase flow: a wetting phase and a non-wetting phase, usually water and air respectively. The following system is considered:

$$\alpha \in \{\text{air}; \text{water}\}, \qquad \begin{cases} \partial_t(\rho_\alpha \Phi S_\alpha) + \nabla \cdot (\rho_\alpha \boldsymbol{q_\alpha}) = 0, \\ \boldsymbol{q_\alpha} = -\dfrac{\mathbb{k}(S_\alpha)}{\mu_\alpha}\nabla(p_\alpha + \rho_\alpha g z), \end{cases} \tag{I.36}$$

where ρ_α denotes the density $[\mathrm{M}\cdot\mathrm{L}^{-3}]$, Φ the porosity [-], S_α the saturation [-], q_α the Darcy velocity $[\mathrm{M}\cdot\mathrm{T}^{-1}]$, $\mathbb{k}$ the permeability tensor $[\mathrm{L}^2]$, μ_α the dynamic viscosity $[\mathrm{M}\cdot\mathrm{L}^{-1}\cdot\mathrm{T}^{-1}]$, p_α the pressure $[\mathrm{M}\cdot\mathrm{L}^{-1}\cdot\mathrm{T}^{-2}]$, t the time [T], z the elevation [L] and g the gravitational acceleration $[\mathrm{L}\cdot\mathrm{T}^{-2}]$. Each of these variables evolve in time inside a three-dimensional space for the general case. The Darcy velocity q_α is modelled by Darcy-Buckingham law as an extension of Darcy's law, initially devoted to single-phase flow, to a two-phase flow system to a two-phase flow system. This modelling implies that both Darcy's law itself and its extension are valid, requiring in particular an assumption of local equilibrium [64].

In the Equation system (I.36), if water contents and relative permeabilities are given, there are four unknowns for two equations: saturations and Darcy velocities for both air and water equations. The system is closed by two equations for saturation and capillary pressure:

$$\begin{cases} S_{\text{air}} + S_{\text{water}} = 1, & \text{by definition,} \\ p_{\text{air}} - p_{\text{water}} = P_{\text{c}}(S_{\text{water}}), \end{cases} \tag{I.37}$$

where P_{c} is capillary pressure $[\text{M} \cdot \text{L}^{-1} \cdot \text{T}^{-2}]$, an invertible function known from experiment.

A major hypothesis for Richards' equation is to eliminate the equation for air in the system (I.36). It requires to assume that the air phase is connected continuously with the atmosphere at every points inside the porous medium. Besides, air viscosity is about 50 times smaller than water viscosity [22]. This means pressure gradients balance much faster in air than in water phase to equilibrate at atmospheric pressure. In other words, air pressure is hydrostatic and thus:

$$\nabla(p_{\text{air}} + \rho_{\text{air}} g z) = 0 \tag{I.38}$$

$$\iff p_{\text{air}} = p_0 - \rho_{\text{air}} g z, \tag{I.39}$$

where p_0 $[\text{M} \cdot \text{L}^{-1} \cdot \text{T}^{-2}]$ is atmospheric pressure which is assumed to be constant at porous medium scale and during the observation time. It is worth mentioning that trapped air pockets and such effects, which do not satisfy the continuous connection assumption for air phase, cannot be treated by the Richards' approach.

Consequently, one can know air pressure everywhere and, if water saturation is known, air saturation is completely defined thanks to the relation (I.37). That is why the continuity equation for water phase from the system (I.36) is now solely considered. Henceforth, the subscript $\square_{\text{water}}$ for water phase is omitted for the sake of clarity. So it writes:

$$\partial_t(\rho \Phi S) - \nabla \cdot \left(\frac{\rho}{\mu} \mathbb{k}(S) \nabla(p + \rho g z) \right) = 0. \tag{I.40}$$

Some reasonable additional hypotheses are usually taken into consideration:

- (H1) water density is sufficiently homogeneous so that $\nabla \rho = 0$;
- (H2) water is sufficiently incompressible so that $\partial_t \rho = 0$;
- (H3) the solid skeleton is sufficiently not deformable so that $\partial_t \Phi = 0$.

Hypotheses (H1) and (H2) lead to rewrite the equation (I.40) such that:

$$\partial_t(\Phi S) - \nabla \cdot \left(\frac{\rho g}{\mu} \mathbb{k}(S) \nabla \left(\frac{p}{\rho g} + z \right) \right) = 0. \tag{I.41}$$

Here, a digression is undertaken to give an understanding about the underlying dynamics of Richards' equation. Knowing that air density is about 820 times smaller than water density at sea level at 15°C, the following quantity can be approximated:

$$\nabla \frac{p}{\rho g} = \nabla \frac{p_{\text{air}} - P_{\text{c}}}{\rho g} = \nabla \frac{p_0 - \rho_{\text{air}} g z - P_{\text{c}}}{\rho g} \approx -\nabla \frac{P_{\text{c}}}{\rho g}. \tag{I.42}$$

This allows some manipulations from equation (I.41) to give:

$$\partial_t(\Phi S) - \frac{\rho g}{\mu} \nabla \cdot (\mathbb{k}(S) \nabla z) + \frac{1}{\mu} \nabla \cdot (\mathbb{k}(S) \nabla P_{\text{c}}) = 0. \tag{I.43}$$

This extended —but somehow approximated —formulation separates the convective term associated with gravitational forces and the diffusive term associated with capillary forces. This shows the complex convective-diffusive dynamics modelled by Richards' equation in spite of neglected air phase effects.

Going back to equation (I.41), water content (θ, [-]), pressure head (ψ, [L]) and hydraulic conductivity tensor ($\mathbb{K}$, $[\mathrm{L}\cdot\mathrm{T}^{-1}]$) are introduced:

$$\theta(S) = \Phi S, \tag{I.44}$$

$$\mathbb{K}(S) = \frac{\rho g}{\mu}\Bbbk(S), \tag{I.45}$$

$$\psi = \frac{p}{\rho g}. \tag{I.46}$$

Water content θ with hypothesis (H3) and hydraulic conductivity $\mathbb{K}$ depend only on the saturation S and not on the primary variable ψ, the pressure head, like it is usually done in Richards' equation. However, the writing is straightforward since the saturation is directly linked to pressure head though the invertible function of capillary pressure P_c:

$$S = P_\mathrm{c}(p) = P_\mathrm{c}(\rho g\psi). \tag{I.47}$$

This relation, often named water retention curve or soil moisture characteristic, is fundamental in study of porous media flow to define hydraulic properties. It will be discussed further in section Section I.4.

Three main formulations of Richards' equation can be established: the pressure-based formulation, the saturation-based formulation or the mixed formulation. The former two formulations contain a single unknown but they are either non-conservative or undefined for complete saturation and heterogeneous soils. The pressure-based formulation and the saturation-based formulation are written respectively as:

$$C(\psi)\partial_t\psi - \nabla\cdot(\mathbb{K}(\psi)\nabla(\psi + z)) = 0, \tag{I.48}$$

$$\partial_t\theta - \nabla\cdot(\mathbb{D}(\theta)\nabla\theta + \mathbb{K}(\theta)\nabla z) = 0, \tag{I.49}$$

with the capillary capacity C $[\mathrm{L}^{-1}]$ and the hydraulic diffusivity $\mathbb{D}$ $[\mathrm{L}^2\cdot\mathrm{T}^{-1}]$ which are respectively defined as:

$$C(\psi) = \frac{\mathrm{d}\theta(\psi)}{\mathrm{d}\psi}, \tag{I.50}$$

$$\mathbb{D}(\theta) = \mathbb{K}(\theta)\frac{\mathrm{d}\psi(\theta)}{\mathrm{d}\theta} = \frac{\mathbb{K}(\theta)}{C(\theta)}. \tag{I.51}$$

In the pressure-based formulation, fluid compressibility effects may be included and then accounting for hypothesis (H2). The usual capillary capacity definition is replaced by the following expression:

$$C(\psi) = \frac{\mathrm{d}\theta(\psi)}{\mathrm{d}\psi} + S_\mathrm{s}S_\mathrm{a}, \tag{I.52}$$

where S_s denotes the specific storage $[\mathrm{L}^{-1}]$ and S_a the aqueous phase saturation [-].

The mixed formulation is selected in the present study owing to its versatility and better numerical behaviour in spite of holding two linked unknowns [23]. The mixed formulation of Richards' equation can finally be written in pressure head form:

$$\partial_t\theta(\psi) - \nabla\cdot(\mathbb{K}(\psi)\nabla(\psi + z)) = 0. \tag{I.53}$$

The mixed formulation can be rewritten in the hydraulic head form which is more common in hydrology:

$$\partial_t \theta(h - z) - \nabla \cdot (\mathbb{K}(h - z)\nabla h) = 0, \tag{I.54}$$

where $h = \psi + z$ is the hydraulic head [L]. One can also add a source/sink term Q [T^{-1}] to Richards' equation to model various processes, like bacteria colony or plant roots, or to couple Richards' equation with free surface flow.

I.4 Hydraulic properties

Solving Richards' equation (I.53) requires two constitutive laws: one for water content and one for hydraulic conductivity. Several models exist, depending on the hydraulic properties of the porous medium.

Usually, the tensor of permeability $\mathbb{k}$, and so the hydraulic conductivity $\mathbb{K}$, are supposed to react to saturation identically for each space direction. This leads to write:

$$\mathbb{k}(\psi) = \mathbb{k}_{\mathrm{s}} k_{\mathrm{r}}(\psi), \tag{I.55}$$
$$\mathbb{K}(\psi) = \mathbb{K}_{\mathrm{s}} K_{\mathrm{r}}(\psi), \tag{I.56}$$

where $\mathbb{k}_{\mathrm{s}}$ denotes the intrinsic or saturated permeability tensor [L^2], k_{r} the relative permeability [-], $\mathbb{K}_{\mathrm{s}}$ the intrinsic or saturated hydraulic conductivity tensor [$\mathrm{L} \cdot \mathrm{T}^{-1}$] and K_{r} the relative hydraulic conductivity [-].

For practical purposes, the water content is often described in terms of effective saturation S_{e} [-]:

$$S_{\mathrm{e}}(\psi) = \frac{\theta(\psi) - \theta_{\mathrm{r}}}{\theta_{\mathrm{s}} - \theta_{\mathrm{r}}}, \tag{I.57}$$

where θ_{s} denotes the saturated water content [-] and θ_{r} the residual water content [-], corresponding to the maximal and minimal saturations, respectively.

The air entry pressure is denoted by ψ_e. It defines saturated and unsaturated zones inside the porous medium. Most of the time, ψ_e is set to 0. The saturated zone of the porous medium corresponds to $\psi > \psi_e$ while the unsaturated or vadose zone corresponds to $\psi < \psi_e$. The water table corresponds to $\psi = \psi_e$ by definition and is considered belonging to the saturated zone. The capillary fringe is the layer above water table where water is raised due to capillary actions. It belongs to the unsaturated zone and there is no standard definition for its upper limit. The hydraulic properties present two different behaviours depending on whether the porous media is saturated ($\psi \geq \psi_e$) or not ($\psi < \psi_e$):

$$S_{\mathrm{e}}(\psi) = \begin{cases} 1 & \text{if } \psi \geq \psi_e, \\ S_{\mathrm{e}} & \text{otherwise,} \end{cases} \quad \text{and} \quad K_{\mathrm{r}}(\psi) = \begin{cases} 1 & \text{if } \psi \geq \psi_e, \\ K_{\mathrm{r}} & \text{otherwise.} \end{cases} \tag{I.58}$$

S_{e} and K_{r} are monotonic increasing functions of pressure head ψ in the unsaturated zone. Throughout the thesis, several constitutive laws will be used to model hydraulic properties in the unsaturated zone. They are compiled in Tab. I.1.

Name	Expression	Parameters
Gardner-Irmay relations (1958) [68, 69]	$S_e = e^{\frac{\alpha\psi}{m}}$ $K_r = e^{\alpha\psi}$	α: pore-size distribution [-] m: tortuosity [-]
Vachaud's relations (1971) [70]	$S_e = \frac{C}{C + \lvert\psi\rvert^D}$ $K_r = \frac{A}{A + \lvert\psi\rvert^B}$	A, C: empirical shape parameters [$L^{B;D}$] B, D: empirical shape parameters [-]
Van Genuchten-Mualem relations (1980) [71, 72]	$S_e = (1 + (\alpha\lvert\psi\rvert)^n)^{-m}$ $K_r = S_e^l \left(1 - \left(1 - S_e^{\frac{1}{m}}\right)^m\right)^2$	$l = \begin{cases} 0.5 & \text{for Mualem [71]} \\ 1 & \text{for Burdine [73]} \end{cases}$: pore connectivity [-] α: parameter linked to air entry pressure inverse [L^{-1}] $n > 1$: pore-size distribution [-] $m = 1 - \frac{1}{n}$: pore-size distribution [-]
Modified Van Genuchten-Mualem relations (2001) [74]	$S_e = (1 + (\alpha\lvert\psi\rvert)^n)^{-m}$ $K_r = \left(\frac{S_e}{S_e^*}\right)^l \left(\frac{1 - \left(1 - S_e^{\frac{1}{m}}\right)^m}{1 - \left(1 - S_e^{*\frac{1}{m}}\right)^m}\right)^2$	$S_e^* = (1 + (-\alpha\psi_s)^n)^{-m}$ [-] with $\psi_s \leq 0$ and $\psi_e = \psi_s$ $l = \begin{cases} 0.5 & \text{for Mualem [71]} \\ 1 & \text{for Burdine [73]} \end{cases}$: pore connectivity [-] α: parameter linked to air entry pressure inverse [L^{-1}] $n > 1$: pore-size distribution [-] $m = 1 - \frac{1}{n}$: pore-size distribution [-]

Table I.1 – Hydraulic relations used in this thesis.

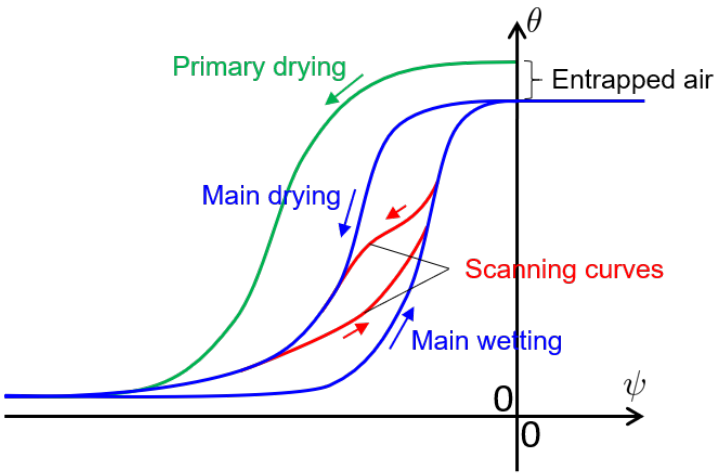


Figure I.16 – Typical soil water characteristic curve with key elements.

A well-known experimental fact is not taken into account by these models: hydraulic properties are hysteretic and dynamic, see Fig. I.16. Indeed, whether water is filling or draining the pores, different wetting (imbibition) and drying (drainage) curves may be distinguished. Besides, dynamics can appear due to memory effects or capillary actions. Equilibrium is located between the two main curves following history of wetting/drying scanning curves. The difference of water content between primary drying and main wetting curves is due to occluded air entrapped by fast infiltration. A similar effect can be observed for primary wetting curve with residual water. Each main curve can be approximated by asymptotics which define cut-off points: the air-entry value (drying) and the air-expulsion value (wetting). They delimit three zones along the curves which accounts for air/water phase continuity: the boundary effect zone (capillary regime), the transition zone (funicular regime) and the residual zone (pendular regime).

Even though the problem of hysteresis is well recognized, these phenomena are often ignored to keep the mathematical description of hydraulic properties simple. In other words, the hydraulic properties are approximated by constitutive laws showing a shape of transfer functions. Actually, hysteresis takes part in few situations, usually at small level, like for viscous fingering. Undergoing works try to integrate effectively these hysteretic and dynamic behaviours into hydraulic properties for Richards' equation [75, 76] but they are still limited and lack assessment.

It is worth noting that:

- under complete saturation, hydraulic properties become constant and Richards' equation degenerates into an elliptic equation characterised by fast diffusion;
- under almost complete unsaturation, hydraulic properties get very near-zero values which stop diffusion and may be inconvenient numerically;
- for particular set of parameters, when $\psi \to 0^-$, constitutive laws may exhibit very steep gradients and even a numerical singularity. Nevertheless, in practice, these gradient-values can be large but are bounded.

These constitutive laws are mainly responsible for the numerical challenges of Richards' equation, such as nonlinearities, degenaracies and instabilities, which are often observed in the presence of material heterogeneities or dynamic boundary conditions like seepage boundary condition.

I.5 Seepage and dynamic forcing boundary conditions

The seepage boundary condition is specific to subsurface model [77]. This condition is used to model the interface between a porous medium and the atmosphere. If the porous medium is saturated and outflow occurs, then water pours out at atmospheric pressure: $\psi = 0 \iff h = z$. Otherwise, the interface acts as an impervious boundary and there is no flux. This condition mimics an outflow condition. Figure I.17 depicts the situation. The treatment of the seepage boundary condition is difficult because the length of seepage face is unknown *a priori*: it depends on the sought solution. Therefore, seepage faces are associated to a dynamic nonlinear boundary condition. Paniconi and Putti [17] reported that seepage condition can cause oscillations leading to convergence problems if rapid changes of saturation state occur near the seepage face. This holds no matter convergence requirements are relaxed or not for the nonlinear iterative solver.

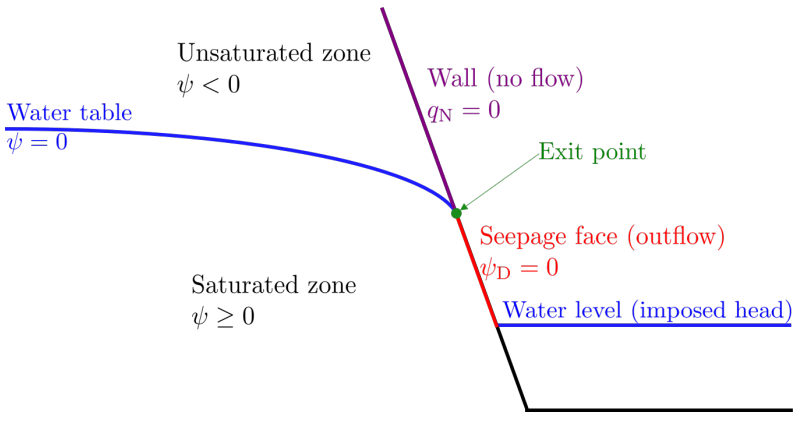


Figure I.17 – Seepage modelling by the boundary condition

There are many ways to numerically express the seepage boundary condition. A simple approach is to consider a switch between a Dirichlet ψ_{D} and a Neumann q_{N} boundary condition:

$$\begin{cases} h = z & \text{if } h \geq z \text{ and } -\mathbb{K}(h-z)\nabla h \cdot \boldsymbol{n} > 0, \\ -\mathbb{K}(h-z)\nabla h \cdot \boldsymbol{n} = 0 & \text{otherwise.} \end{cases} \tag{I.59}$$

Another way to take into account the seepage boundary condition in many works [78] is to write it as a Signorini's type:

$$\psi \leq 0, \quad -\mathbb{K}(h-z)\nabla h \cdot \boldsymbol{n} \geq 0, \qquad (-\mathbb{K}(h-z)\nabla h \cdot \boldsymbol{n})\psi = 0. \tag{I.60}$$

Alternatively, the seepage boundary condition may be interpreted as a nonlinear Robin boundary condition:

$$\mathbb{1}_{\mathrm{S}}(h)(h-z) - (1-\mathbb{1}_{\mathrm{S}}(h))\mathbb{K}(h-z)\nabla h \cdot \boldsymbol{n} = 0, \tag{I.61}$$

where the seepage indicator function lying on the boundary Γ_{S} is:

$$\begin{aligned} &\mathbb{1}_{\mathrm{S}} \colon \Gamma_{\mathrm{S}} \to \{0,1\}, \\ &\qquad h \mapsto \begin{cases} 1 & \text{if } h \geq z \text{ and } -\mathbb{K}(h-z)\nabla h \cdot \boldsymbol{n} > 0, \\ 0 & \text{otherwise.} \end{cases} \end{aligned} \tag{I.62}$$

This point of view makes its writing compact and that is why it will be used later for the discretization, keeping in mind that it encompasses a mix of a Dirichlet and a Neumann boundary conditions such as in (I.59).

For this thesis, a dynamic forcing boundary condition will be used. It allows to monitor a set of two conditions: one is a Dirichlet boundary condition and the other is a seepage boundary condition. The Dirichlet boundary condition for hydraulic head is both space- and time-dependant which makes it very dynamic. The seepage boundary condition is applied above the Dirichlet boundary condition, over the remaining boundary. This dynamic forcing boundary condition models moving water outside the seeping porous medium which may infiltrate at the prescribed hydraulic head. Phenomenon like reservoir filling, flooding, ponding or waves can be approximated with this kind of boundary condition.

I.6 Numerical challenges of Richards' equation

I.6.1 Behaviour analysis

The solution of Richards' equation holds typical features of groundwater flows. The key process concerns water moving through the unsaturated porous medium under the combined actions of gravity and capillarity.

When infiltration occurs, the interface between the initial state of soil and the newly wetted zone is called **wetting front**. The wetting front is moving dynamically and can be extremely sharp according to hydraulic properties and variable saturation conditions. It is linked to the nonlinearity varying diffusivity of hydraulic properties which control its steepness. Similar behaviour may be encountered for a **drying front**.

Another flow phenomenon is associated with porous medium heterogeneity and anisotropy. The interface between two materials of different hydraulic properties can lead to rapid change of saturation

states called **internal layer**. This internal layer is static since it is attached to the material interface. It can be extremely sharp and is often linked to a discontinuity, in particular for water content. The jump depends to the saturation conditions and hydraulic properties on both sides. When a wetting front hits an internal layer, capillary forces can prevent water to infiltrate further. Water accumulates until it is able to break through. This is called the **capillary barrier** effect.

One may also want to connect the porous medium with another domain. For example, this is quite natural when a coupling with free surface water is considered though Beavers-Joseph-Saffman boundary condition. Then, the interface between the two domains should be checked carefully since it is related to some discontinuities and **boundary layer** phenomenon.

Spurious oscillations in the solution of Richards' equation are a non-physical effect and should not be present. They are found for the aforementioned flow patterns. These oscillations consist typically in undershoots/overshoots ahead of steep wetting fronts which are moving in an initially very dry zone.

Here, a brief qualitative analysis and some suggestions are outlined to understand the behaviour of Richards' equation. Richards' equation is classified as a quasilinear degenerate parabolic equation. It is quasilinear because nonlinearity depends only on the solution. It is degenerate parabolic because under complete saturation, hydraulic properties are constant so that Richards' equation becomes elliptic. Nevertheless, Richards' equation shares properties with convection-diffusion equation and can behave like them. For example, sharp wetting fronts and internal layers can occur which leads to instabilities in the form of spurious oscillations. This not so surprising if Richards' equation is considered as an approximation from a diphasic system of convection-diffusion equation [79, 80]. Richards' equation is mainly diffusive as an elliptic-parabolic equation but convection and diffusion are locally in competition around these sharp fronts and layers. For this reason, making analogy between Richards' equation and heat equation is useful [81] but it is limited.

Richards' equation is predominantly diffusive. In saturated zones, it is elliptic and so purely diffusive: the time derivative is zero and there is no advection. Source terms and boundary conditions impose the solution state. In unsaturated zones, two states may be distinguished considering Eq. (I.43) which is a form of Fokker-Planck equation. Richards' equation is parabolic and mainly diffusive in unsaturated zone but Richards' equation may exhibit an hyperbolic behaviour where advection is dominant. This latter state occurs typically for sharp wetting fronts. In practice, the behaviours of Richards' equation are often mixed because hydraulic properties governing flow patterns are nonlinear. Then, the study and prediction of the dynamic response of the equation is difficult. Some attempts are found in literature to analyse Richards' equation fronts, for instance in [82, 83].

It is hard to represent Richards' equation locally by a linear approximation that is why Richards' equation is sometimes referred as highly nonlinear. Moreover, Richards' equation is known to be a stiff differential equation because some numerical schemes are numerically unstable unless the time step-size is taken to be extremely small. Stiffness is difficult to define but means that the equation holds some terms leading to rapid variation in the solution.

I.6.2 Computational developments

The oscillations arise whatever the numerical method is used: finite difference, finite volume, finite element or discontinuous Galerkin methods. However, they are seldom observed in finite differences because a suitable scheme prevents them like upwinding. In finite volume and discontinuous Galerkin methods, tools like slope limiters make oscillation-free solution possible [28, 84]. These techniques may delay the wetting front and are not straightforward for multidimensional or general mesh because

they depend on cells geometry. For finite elements, it is more difficult to eliminate the non-physical oscillations. Mass lumping [85–88] or mass distributed [86] schemes are usually employed because it was seen that they reduce oscillations but they may smear the wetting front in return. Additional techniques are often necessary to stabilize the solution: basis function respecting the maximum principle, multiscale formulation [80], velocity field postprocessing [80], *etc*.

The origin of these spurious oscillations is not well identified in literature since many explanations and analysis are given: saturation distribution, velocity field regularity, discrete hydraulic conductivity scheme [89], M-matrix property [88]. Notably, refinement always eliminates the spurious oscillations around the wetting front and makes the numerical solution converge to the physical solution. This observation makes think that spatial resolution is not sufficient to capture the sharpness of the front. Even though the solution is mathematically regular enough, the numerical solution is unstable and can even be discontinuous because of phenomena occurring at scale smaller than elements size [90]. Global refinement is not the answer because the needed element size is so small that the computing cost is irrelevant [80]. That is why local and dynamic refinement is investigated for Richards' equation. The development of such techniques is demanding because it must capture moving wetting fronts and add enough elements.

A useful property for quality assessment is the monotonicity preservation guaranteed by the maximum principle. Indeed, hydraulic conductivity and water content are bounded and (positive) monotonic increasing functions of pressure head. So, in the absence of source/sink terms, the solution should be locally bounded according to the boundary conditions [79, 88, 91]. This is known as maximum principle. A monotonic solution respects the maximum principle which can serve for instance to detect non-physical solutions holding overshoots/undershoots.

Summary

Richards' equation is largely used to model flows in unsaturated porous media. It is very attractive since it allows to model the porous medium as a whole part, that is to say both the saturated and unsaturated zones, without actually using a two-phase system. However, some limitations can be drawn. For instance, the porous medium is considered incompressible and water is incompressible and homogeneous but theses assumptions can be removed with extended formulations of Richards' equation. Hydraulics properties of Richards' equation are modelled in an equilibrium state as well since they lack dynamic hysteresis effects. Recent developments of Richards' equation attempt to incorporate such dynamics to take into account capillary behaviour. Stronger assumptions imply that the porous medium is inert and water is isotherm. Moreover, Richards' equation lies on modelling hypotheses. Firstly, upscaling from microscale Navier-Stokes equation to macroscale Darcy-Buckingham law, which the basis relation for Richards' equation, shows flow has to be sufficiently slow with small enough pores to neglect inertia terms. Secondly, air-phase effects, like trapped air pockets or bubbles, are completely neglected since air phase has to be at atmospheric pressure and continuously connected to be disregarded.

Richards' equation is a parabolic equation, predominantly diffusive. It has been employed for a long time so community has gathered a lot of experience about its numerical solution. However, it can degenerate into an elliptic equation and holds nonlinear hydraulic properties. Then, Richards' equation exhibits different dynamics depending on the saturation state which makes its solution troublesome. The different formulations of Richards' equation involve different set of variables and are not equivalent from a numerical point of view. The mixed formulation has better properties because it is mass conservative and handle automatically saturated-unsaturated zones. Mass conservation is an important property to reduce mass balance errors and model properly water fluxes. One of the main numerical issue concerns moving sharp wetting fronts. They produce unphysical oscillations in the solution. A large number of techniques have been used to deal with them, sometimes in combination, but these approximations lack inter-comparison and assessments. So, it is difficult to know whether or not such techniques is sufficient or necessary. To contribute to the analysis of Richards' equation, its solution has been interpreted within the context of convection-diffusion equation. For this purpose, behaviour is sorted between pure, vanishing, heterogeneous, anisotropic and nonlinear varying diffusivity.

Chapter II

Numerical methods for Richards' equation

"Tú, que me lees, ¿estás seguro de entender mi lenguaje?"
— Jorge Luis Borges, *The Library of Babel*, 1941

The numerical method used to solve Richards' equation is described in this chapter which can be formally divided into two parts.

The first part deals with the discretization of Richards' equation both in space and time. Provided numerical issues discussed in the previous chapter, the discretization of Richards' equation is still a matter of interest in order to get accurate and efficient results for a large range of problems which is pointed out by a short review. The space discretization method should hold a certain number of features to be interesting compared to numerous existing discretization for Richards' equation. In this work, discontinuous Galerkin (DG) methods are chosen because they provide flexibility in the numerical scheme and show promising developments. For instance, high-order approximation, adaptive mesh refinement or local mas balance are natural abilities of DG methods. For time discretization, implicit scheme are favoured for Richards' equation because they are stable for any time step. Backward difference formulas are chosen for the time dizcretization because they are known to be good for stiff partial differential equations and can reach high-order with less cost than other methods like Runge-Kutta methods.

The second part is about the treatment of nonlinearities from Richards' equation. In particular, the nonlinear solver based on the linearization of Richards' equation by the fixed-point iteration and Newton-Raphson methods is presented. The iterative procedure comes with adaptive time stepping in order to give reliability to the nonlinear convergence.

II.1 A brief review

II.1.1 General overview of discretization methods

Lots of discretization techniques were employed for Richards' equation with the use of finite difference (FD) methods until the 1980s. Here, some examples are listed. FD methods were used by Celia and Bouloutas [85] or Woodward and Dawson [92] or Dogan and Motz [93]. Finite volume (FV) methods were used by Manzini and Ferraris [94], Orgogozo *et al.* [95] or Caviedes-Voullième *et al.* [96]. Finite element (FE) methods are found in Celia and Bouloutas [85], Lehmann and Ackerer [18], Šimůnek [97], Forsyth *et al.* [91], Katvetsi *et al.* [98]. Diersch and Perrochet [99], Knabner and Schneid [100], Chavent and Roberts [101] or Fahs *et al.* [102] used mixed FE methods. Others techniques are more sparsely found like a multiscale FE methods [103] and a lattice Boltzmann method [104]. More information is available in the reviews from Farthing *et al.* [23] and Zha *et al.* [24].

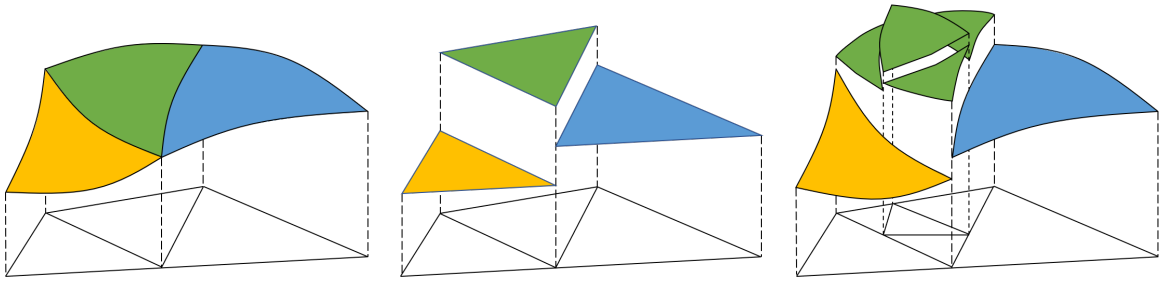
Concerning time discretization, the common approach when working with Richards' equation is to use implicit schemes to avoid the Courant–Friedrichs–Lewy (CFL) condition which is very restrictive for some hydraulic properties. Implicit Euler or Crank-Nicolson schemes are widespread for solving Richards' equation. An ongoing trend is to use high-order accuracy in time with, for example, backward differentiation formula (BDF) [78, 105]. In this study, the implicit Euler scheme and BDF methods will be used because they are suitable for stiff equations like Richards' equation. The first one is simple and shows a wide region of stability while the others give time high-order to balance space and time discretization errors. More information is available in the reviews from Farthing *et al.* [23] and Zha *et al.* [24].

Two nonlinear iterative solvers are presented in this study to solve the algebraic system of equations arisen from discretization: a damped Newton-Raphson method and a fixed-point method (or Picard's method). Previous studies assessed nonlinear behaviour of Richards' equation through various cases

and multiple numerical schemes [17, 18, 30] in which other nonlinear solving techniques are also presented. Despite quadratic convergence for Newton-Raphson method, there is no clear advantage compared to fixed-point method for a certain number of cases, mainly because Newton-Raphson method often fails to converge due to poor initial estimate, derivative issues or solution overshoots. Fixed-point method appears as more robust. Moreover, time adaptation can help managing Richards' equation solution and relax the nonlinear iterative solver.

II.1.2 Principles of discontinuous Galerkin methods

DG methods are relatively recent in history of discretization methods since they date from 1973 when Reed and Hill introduced them in the context of neutron transport problem [106]. They might be seen as a generalization of the FE or the FV framework. Indeed, they share properties from both of these methods which makes them attractive. Figure II.1 shows the typical differences of solution approximation between FE methods, FV methods and DG methods. FE methods are continuous approximation while FV methods are discontinuous approximation. In general, high-order approximation is doable with FE methods but much more difficult with FV methods. DG methods have high-order discontinuous approximation making mesh adaptation very easy.



(a) Finite element methods (b) Finite volume methods (c) Discontinuous Galerkin methods with mesh refinement

Figure II.1 – Typical example of a 2D-solution representation arising from different methods of approximation.

More precisely, DG methods have the following advantages:

- They give rise to an algebraic system of decoupled equations, that is to say matrices in the discrete system are sparse and structured by blocks. Each block has a size equals to the local number of degrees of freedom (belonging to one element). Simpler solving techniques can be employed and parallelization can be achieved;
- Stencil size is independent of the order of approximation. Then, solving on one element depends only on the solution of this element and its direct neighbours. The method is said to be compact. High-order approximation is easier to reach.
- Numerical approximation of the solution is only locally smooth that is to say over one element. First, discontinuity can be approximated like FV methods. Besides, solution sensibility to mesh regularity can be alleviated through the discontinuity between elements. In particular, this allows to work easily with meshes which are unstructured, non-conforming and hybrid (several elements geometry) giving path to mesh refinement;

- Boundary conditions are not imposed through the definition of approximation space but rather directly incorporated in the weak formulation. Study and implementation are then easier;
- DG methods are local making them very suitable for conservation laws. Numerical dissipation is low and they are locally mass conservative.

DG methods are very compliant methods handling many desirable features.

Yet, DG methods also have disadvantages:

- DG numerical solution can hold spurious oscillations which are not physical. This is due to numerical stability ensured by penalty or numerical dispersion at discontinuity. In this sense, DG methods lack robustness. Lots of techniques are developed to deal with these oscillations like slope limiters, artificial viscosity, penalty definition, numerical flux in the numerical scheme;
- The computational cost can be prohibitive because the number of degrees of freedom is multiplied compared to the standard FE method. For DG methods, the number of degrees of freedom is proportional to the number of elements depending on approximation order while, for FE methods, the number of degrees of freedom is linked to the number of vertices (and possibly elements);
- Implementation for DG methods is less spread than other discretization techniques due to its relatively recent use. Then, it can be harder to achieve for new-comers. Besides, comparatively to FE methods, degrees of freedom may have no physical meaning which may trouble new users;
- Even if DG methods analysis has made great progress, optimal error estimate are not fully derived and there are still gaps between theoretical rates of convergence and observed rates.

DG methods are more competitive nowadays since parallelization and computers are becoming more suited to have great computation and memory capacities. Besides, approximation can be better than FE methods but hold an equivalent number of degrees of freedom. For instance, DG methods can have less but higher order elements [27] or use approximation space less rich than the ones needed in FE methods [26]. One can also remark that numerical dispersion in DG methods stays locally restrained compared to FE methods [27].

Further developments extended the use of DG methods to other study fields with elliptic and parabolic equations and new forms appeared such as primal and local formulations [107]. Despite this history of developments and their overall properties, DG methods were sparsely applied to Richards' equation up to now. Here, such studies are listed as exhaustively as possible. In 2007, Li *et al*. solved Richards' equation in 1D with a local DG method [21, 31] which is a mixed formulation. Sochala discretized Richards' equation through a SIPG mixed DG formulation in 2008 [78, 84]. Krämer used a mixed DG formulation during his master thesis work in 2016 [76]. Recently, in 2019, Dolejší *et al*. proposed a space-time DG method for solving Richards' equation [108]. Nevertheless, it should be noticed that DG methods are more and more used for complex flow and transport phenomena in porous media described by other models such as two-phase flow problems [28, 29, 109]. In this study, Richards' equation is approximated by primal interior penalty DG methods as matter of interest but also because this formulation is simpler.

Finally, making the most of DG methods, *hp*-adaptivity can improve solution quality of Richards' equation. So far, *p*-adaptation is not really adopted on its own for flows problems in porous media; whereas *h*-adaptation is employed largely for two-phase flow in porous media [28] and Richards' equation [20, 31]; *hp*-adaptation has been exploited in many fluid dynamics applications but few works use it for porous media flow [110, 111]. As far as author knows, for Richards' equation, three

studies are concerned: Li *et al.* propose an *hp*-local DG approximation [21]; Šolín and Kuraz use an *hp*-FE method [112]; Dolejší *et al.* present an anisotropic *hp*-adaptive space-time DG method [108]. These *hp*-adaptative algorithm are not equivalent since they aim for different goals, rely on distinct approaches to make decisions and use different techniques to refine/enrich the mesh/polynomial order. In general, such questions remain open problems for *hp*-adaptation [113].

II.2 Space semidiscretization

Many DG methods [107] exist but they can be viewed according to two approaches: the primal formulation and the mixed formulation. The former corresponds to a model problem written as one single equation. The latter corresponds to a model problem broken down into a potential and flux parts so that it forms a system of two coupled equations which are approximated by two functional spaces linked by numerical flux. In this work, the Richards' equation is discretized by several DG methods. These methods are very similar to each other and can be described by the same primal weak formulation. In the following derivation of the variational formulation, all of them are considered. First of all, some notations and definitions are established for the spatial and temporal discretizations in order to get the numerical scheme. More careful developments can be found in Rivière [26] or Dolejší and Feistauer [27].

II.2.1 Fundamental concepts

Problem statement and notations

Let $d \in \{1, 2, 3\}$ be the dimension. The porous medium is represented by the computational domain $\Omega \subset \mathbb{R}^d$ of boundary $\partial\Omega$ and the final time is $T \in \mathbb{R}_+^*$. The boundary $\partial\Omega$ is subdivided into three mutually disjoint boundaries, $\partial\Omega = \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{N}} \cup \Gamma_{\mathrm{S}}$, corresponding to the Dirichlet, Neumann and seepage boundary conditions respectively. The following problem is considered:

find $h(\boldsymbol{x}, t) : \Omega \times (0, T) \longrightarrow \mathbb{R}$ such that

$$\begin{cases} \partial_t \theta(h - z) - \nabla \cdot (\mathbb{K}(h - z)\nabla h) = 0, & \text{in } \Omega \times (0, T), \\ h = h_0, & \text{in } \Omega \times \{0\}, \\ h = h_{\mathrm{D}}, & \text{on } \Gamma_{\mathrm{D}} \times (0, T), \\ -\mathbb{K}(h - z)\nabla h \cdot \boldsymbol{n} = q_{\mathrm{N}}, & \text{on } \Gamma_{\mathrm{N}} \times (0, T), \\ \mathbb{1}_{\mathrm{S}}(h)(h - z) - (1 - \mathbb{1}_{\mathrm{S}}(h))\mathbb{K}(h - z)\nabla h \cdot \boldsymbol{n} = 0, & \text{on } \Gamma_{\mathrm{S}} \times (0, T). \end{cases} \tag{II.1}$$

The strong problem (II.1) has a solution $h \in L^2(\Omega \times (0, T))$. The dependence of space and time is omitted for clarity. Let $h_0 \in L^2(\Omega)$, $h_D \in L^2\Big(L^2(\Gamma_D); 0, T\Big)$, $q_N \in L^2\Big(L^2(\Gamma_N); 0, T\Big)$. $\mathbb{K}$ is a monotonically increasing matrix-valued function of h: $\mathbb{K} : \mathbb{R} \longrightarrow [\mathbb{R}_+]^{d\times d}$ which is symmetric, positive definite, bounded below and above uniformly. θ is a monotonically increasing function of h: $\theta : \mathbb{R} \longrightarrow \mathbb{R}_+$ which is bounded below and above uniformly. $\mathbb{K}$ and θ are continuous functions of h for a given porous medium but can hold discontinuities at interface of heterogeneous materials. The existence, the uniqueness and the regularity for the solution of Richards' equation can be found in the work from Alt and Luckhaus [114], Otto [115, 116] or Schweizer [75, 117, 118].

Geometry partition and definitions

The time duration $(0, T)$ is subdivided into N time intervals such that $0 = t^0 < t^1 < \cdots < t^{\mathrm{N}} = T$. Let $n \in \mathbb{N}, 0 \leq n < N$. If the time interval $T^{n+1} = [t^n, t^{n+1}]$ is considered, the corresponding time

step is $\tau^{n+1} = t^{n+1} - t^n$.

Ω is subdivided into N_E mutually disjoint closed polygonal elements E forming a tessellation or mesh. The mesh at the time partition T^n is denoted by $\mathcal{E}_h^n = \{E\}_{E\in\Omega}$. So, $\overline{\Omega} = \bigcup_{E\in\mathcal{E}_h^n} E$. Moreover, one element $E \in \mathcal{E}_h^n$ has a boundary denoted by ∂E, a roundness ρ_E which is the diameter of the largest ball inscribed into E as well as a diameter h_E defined by:

$$h_E = \operatorname{diam}(E) \coloneqq \sup_{\boldsymbol{x},\boldsymbol{y}\in E} |\boldsymbol{x} - \boldsymbol{y}|. \tag{II.2}$$

In 2D, the diameter of a triangle is its largest edge and the diameter of a quadrangle is its largest diagonal or its largest edge. The aspect ratio $\sigma_E = \dfrac{h_E}{\rho_E}$ is a measure of element non-degeneracy and defines the mesh regularity. The geometry definitions and notations are showed in Fig. II.2. For every mesh $\mathcal{E}_h^n$, the space step is set as:

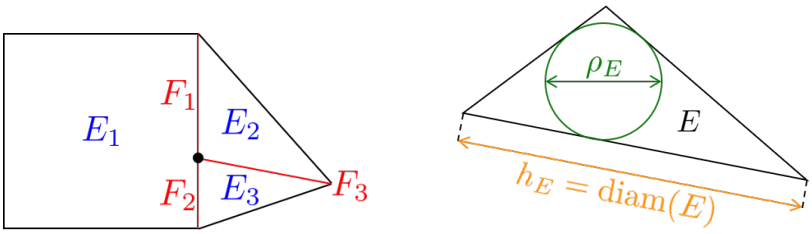
$$h^n \coloneqq \max_{E\in\mathcal{E}_h^n}(h_E). \tag{II.3}$$

A face is defined as the intersection between exactly two elements or between one element and the boundary $\partial\Omega$. Figure II.2 shows a particular example. The set of all open faces of all elements $E \in \mathcal{E}_h^n$ is denoted by $\mathcal{F}_h^n$. Furthermore, these definitions stand:

$$\mathcal{F}_{h,\mathrm{D}}^n \coloneqq \bigcup_{F\in\Gamma_\mathrm{D}} F, \quad \mathcal{F}_{h,\mathrm{N}}^n \coloneqq \bigcup_{F\in\Gamma_\mathrm{N}} F, \quad \mathcal{F}_{h,\mathrm{S}}^n \coloneqq \bigcup_{F\in\Gamma_\mathrm{S}} F, \tag{II.4}$$

$$\mathcal{F}_{h,\mathrm{B}}^n \coloneqq \mathcal{F}_{h,\mathrm{D}}^n \cup \mathcal{F}_{h,\mathrm{N}}^n \cup \mathcal{F}_{h,\mathrm{S}}^n, \quad \mathcal{F}_{h,\mathrm{I}}^n \coloneqq \mathcal{F}_h^n \setminus \mathcal{F}_{h,\mathrm{B}}^n. \tag{II.5}$$

For each face $F \in \mathcal{F}_h^n$, there is a unit normal vector $\boldsymbol{n_F}$. For $F \in \mathcal{F}_{h,\mathrm{I}}^n$, its orientation is arbitrary but kept fixed and, for $F \in \mathcal{F}_{h,B}^n$, it is oriented outward.



(a) Non-conforming mesh with one hanging node

(b) Characteristic dimensions of one element

Figure II.2 – Examples of mesh geometry for discontinuous Galerkin methods.

Functional spaces and operators

Let $s \in \mathbb{N}_+$ and $1 \leq p \leq \infty$. The Sobolev space $W^{s,p}$ is introduced:

$$W^{s,p}(\Omega) = \{v \in L^p(\Omega) : \forall |\alpha| \leq s, D^\alpha v \in L^p(\Omega)\}, \tag{II.6}$$

where α is a multi-index, $D^\alpha v$ is a partial derivative in the sense of distribution and L^p is a Lebesgue space. In the case $p = 2$, the Sobolev space is also a Hilbert space: $H^s(\Omega) = W^{s,2}(\Omega)$. Sobolev spaces may have fractional (real) indices s obtained by interpolating the surrounding Sobolev spaces of

integer indices. A theorem result states that a continuous representative is possible under the following conditions:

$$H^s(\Omega) \subset \mathcal{C}^0(\Omega) \text{ if } \begin{cases} s > \dfrac{1}{2} \text{ for } d = 1, \\ s > 1 \text{ for } d = 2, \quad s > \dfrac{3}{2} \text{ for } d = 3. \end{cases} \tag{II.7}$$

DG methods use discontinuous approximations over the mesh $\mathcal{E}_h^n$ defining a so-called broken Sobolev space $H^s(\mathcal{E}_h^n) \supset H^s(\Omega)$. $H^s(\mathcal{E}_h^n)$ denotes the space of functions whose restriction to each element E belongs to the Sobolev space $H^s(E)$:

$$H^s(\mathcal{E}_h^n) = \left\{ v \in L^2(\Omega) : v\Big|_E \in H^s(E), \forall E \in \mathcal{E}_h^n \right\}. \tag{II.8}$$

The solution is sought in a subspace of this broken Sobolev space, taken to be:

$$S_p(\mathcal{E}_h^n) \coloneqq \left\{ v \in L^2(\Omega) : v\Big|_E \in \mathbb{P}_p(E), \forall E \in \mathcal{E}_h^n \right\}, \tag{II.9}$$

where $\mathbb{P}_p(E)$ denotes the space of polynomial functions on E of total degree less than or equal to $p \in \mathbb{N}$. $\mathbb{P}_p(E)$ can be replaced by $\mathbb{Q}_p(E)$ denoting degree less than or equal to $p \in \mathbb{N}$.

Let two neighbouring elements E_l and E_r sharing one face $F \in \mathcal{F}_{h,\mathrm{I}}^n$. Notations are illustrated on Fig. II.3. There are two traces of a function u along F denoted by u_l and u_r:

$$\boldsymbol{x} \in F, u_\mathrm{l}(\boldsymbol{x}) \coloneqq \lim_{\varepsilon \to 0^-} u(\boldsymbol{x} + \varepsilon \boldsymbol{n_F}), \tag{II.10}$$

$$\boldsymbol{x} \in F, u_\mathrm{r}(\boldsymbol{x}) \coloneqq \lim_{\varepsilon \to 0^+} u(\boldsymbol{x} + \varepsilon \boldsymbol{n_F}). \tag{II.11}$$

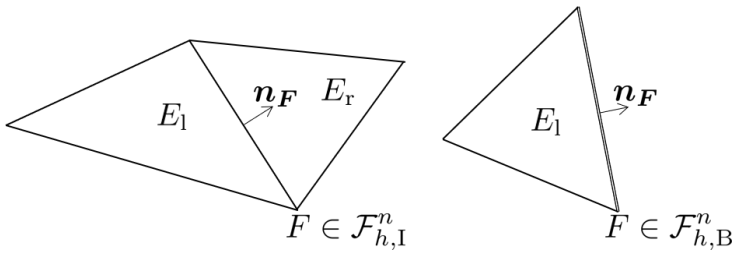


Figure II.3 – Notations for two elements E_l and E_r sharing one face $F \in \mathcal{F}_{h,\mathrm{I}}^n$ and one element E_l with a boundary face $F \in \mathcal{F}_{h,\mathrm{D}}^n$ in 2D.

Formally, the jump and the mean of the function u across a face $F \in \mathcal{F}_{h,\mathrm{I}}^n$ are defined respectively by:

$$[\![u]\!] \coloneqq u_\mathrm{l} - u_\mathrm{r}, \tag{II.12}$$

$$\{\!\{u\}\!\} \coloneqq \frac{1}{2}(u_\mathrm{l} + u_\mathrm{r}). \tag{II.13}$$

The principle is sketched in Fig. II.4. In the case that $F \in \mathcal{F}_{h,\mathrm{B}}^n$, $[\![u]\!] = \{\!\{u\}\!\} = u_\mathrm{l}$ where u_l is the trace of the function u from the element which holds F.

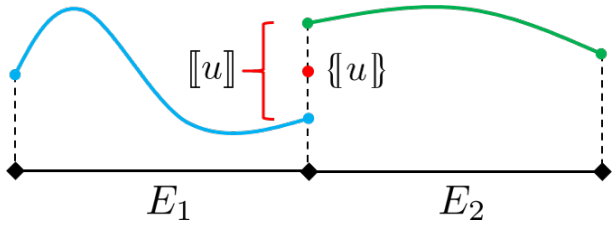


Figure II.4 – Definitions of the mean and jump operators for two elements E_1 and E_2 in 1D.

Throughout this study, the quantity h_F denotes the measure of $F \in \mathcal{F}_h^n$ and p_F:

$$h_F \coloneqq |F|, \tag{II.14}$$

$$p_F \coloneqq \sqrt{\{\!\{p^2\}\!\}}. \tag{II.15}$$

II.2.2 Discontinuous Galerkin weak formulation

The problem (II.1) is multiplied by a test function $v \in S_p(\mathcal{E}_h^n)$ and then integrated on each element $E \in \mathcal{E}_h^n$. After using Green's theorem element-wise, it yields:

$$\begin{cases} \displaystyle\int_E \partial_t\theta(h-z)v\,\mathrm{d}E + \int_E \mathbb{K}(h-z)\nabla h\cdot\nabla v\,\mathrm{d}E - \sum_{F\in\partial E}\int_F \mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F = 0, \\ \displaystyle\int_E hv\,\mathrm{d}E = \int_E h_0 v\,\mathrm{d}E, \\ \displaystyle\sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{D}}}\int_F hv\,\mathrm{d}F = \sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{D}}}\int_F h_{\mathrm{D}}v\,\mathrm{d}F, \\ \displaystyle-\sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{N}}}\int_F \mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n}v\,\mathrm{d}F = \sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{N}}}\int_F q_{\mathrm{N}}v\,\mathrm{d}F, \\ \displaystyle\sum_{F\in\partial E\cap\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)(h-z)v\,\mathrm{d}F - \sum_{F\in\partial E\cap\mathcal{F}_h^{S}}\int_F (1-\mathbb{1}_{\mathrm{S}}(h))\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n}v\,\mathrm{d}F = 0. \end{cases} \tag{II.16}$$

Summing over all elements in $\mathcal{E}_h^n$, it gives:

$$\begin{cases} \displaystyle\sum_{E\in\mathcal{E}_h}\int_E \partial_t\theta(h-z)v\,\mathrm{d}E + \sum_{E\in\mathcal{E}_h}\int_E \mathbb{K}(h-z)\nabla h\cdot\nabla v\,\mathrm{d}E - \sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F [\![\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v]\!]\,\mathrm{d}F \\ \quad\displaystyle-\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F + \sum_{F\in\mathcal{F}_h^{\mathrm{N}}}\int_F q_{\mathrm{N}}v\,\mathrm{d}F \\ \quad\displaystyle-\sum_{F\in\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F = 0, \\ \displaystyle\sum_{E\in\mathcal{E}_h}\int_E hv\,\mathrm{d}E = \sum_{E\in\mathcal{E}_h}\int_E h_0 v\,\mathrm{d}E, \\ \displaystyle\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F hv\,\mathrm{d}F = \sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F h_{\mathrm{D}}v\,\mathrm{d}F, \\ \displaystyle\sum_{F\in\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)(h-z)v\,\mathrm{d}F = 0. \end{cases} \tag{II.17}$$

The Neumann boundary condition has appeared naturally within the formulation and has been substituted. So it stands for the homogeneous Neumann part of the seepage boundary condition.

Assuming that

$$[\![\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}]\!]=0, \tag{II.18}$$

because this quantity is smooth enough, and using the identity:

$$\forall u,v\in S_p(\mathcal{E}_h^n),\ [\![uv]\!]=\{\!\{u\}\!\}[\![v]\!]+\{\!\{v\}\!\}[\![u]\!], \tag{II.19}$$

it yields to:

$$\begin{cases}\displaystyle\sum_{E\in\mathcal{E}_h}\int_E\partial_t\theta(h-z)v\,\mathrm{d}E+\sum_{E\in\mathcal{E}_h}\int_E\mathbb{K}(h-z)\nabla h\cdot\nabla v\,\mathrm{d}E\\ \quad-\displaystyle\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F\{\!\{\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}\}\!\}[\![v]\!]\,\mathrm{d}F-\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F\\ \quad+\displaystyle\sum_{F\in\mathcal{F}_h^{\mathrm{N}}}\int_F q_{\mathrm{N}}v\,\mathrm{d}F-\sum_{F\in\mathcal{F}_h^{S}}\int_F\mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F=0,\\ \displaystyle\sum_{E\in\mathcal{E}_h}\int_E hv\,\mathrm{d}E=\sum_{E\in\mathcal{E}_h}\int_E h_0v\,\mathrm{d}E,\\ \displaystyle\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F hv\,\mathrm{d}F=\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F h_{\mathrm{D}}v\,\mathrm{d}F,\\ \displaystyle\sum_{F\in\mathcal{F}_h^{S}}\int_F\mathbb{1}_{\mathrm{S}}(h)(h-z)v\,\mathrm{d}F=0.\end{cases} \tag{II.20}$$

II.2.3 Penalization

Now, two penalty terms are required to mimic the solution continuity by constraining the interior solution jump to be zero, and to enforce the Dirichlet boundary condition, including the remaining part of the seepage boundary condition. If the solution is a continuous function satisfying the Dirichlet boundary conditions, the two penalty terms are vanishing so they can be added to the formulation:

$$J_{\mathrm{I}}(h,v)=\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F\varrho_F^{\mathrm{I}}[\![h]\!][\![v]\!]\,\mathrm{d}F, \tag{II.21}$$

$$J_{\mathrm{D}}(h,v)=\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F\varrho_F^{\mathrm{D}}hv\,\mathrm{d}F. \tag{II.22}$$

The interior and Dirichlet penalty weights are respectively set as:

$$\varrho_F^{\mathrm{I}}\coloneqq\frac{\sigma_F^{\mathrm{I}}\gamma_F}{\mu_F}, \tag{II.23}$$

$$\varrho_F^{\mathrm{D}}\coloneqq\frac{\sigma_F^{\mathrm{D}}\gamma_F}{\mu_F}, \tag{II.24}$$

where the face measure is defined as $\mu_F\coloneqq\dfrac{{h_F}^{\beta}}{{p_F}^2}$ and where σ_F^{I}, σ_F^{D}, β and γ_F are suitable chosen positive constants. σ_F^{I}, σ_F^{D} are a user-defined parameter which can be seen as a penalty speed. They must be above a threshold value to provide coercivity and guarantee stability for some of the DG

methods [119]. At the same time, it should not be too large since the condition number increases with the penalty parameter leading to instabilities. This was studied by Castillo [120] and for layered reservoirs with high permeability contrast by Slingerland and Vuik [121]. β should be one for standard methods but can take greater values to enforce continuity as in FE method but in response can also degrade numerical properties such as condition number. It can be seen as a penalty order and is referred as superpenalization in the literature. Finally, γ is a diffusion penalty coefficient set to one generally whereas some studies choose other definitions, see *e.g.* [122, 123].

Then, a first weak formulation rises:

$$\begin{cases} \displaystyle\sum_{E\in\mathcal{E}_h}\int_E \partial_t\theta(h-z)v\,\mathrm{d}E + \sum_{E\in\mathcal{E}_h}\int_E \mathbb{K}(h-z)\nabla h\cdot\nabla v\,\mathrm{d}E \\ \quad - \displaystyle\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F \{\!\!\{\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}\}\!\!\}[\![v]\!]\,\mathrm{d}F - \sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F \\ \quad + \displaystyle\sum_{F\in\mathcal{F}_h^{\mathrm{N}}}\int_F q_{\mathrm{N}}v\,\mathrm{d}F - \sum_{F\in\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F + \sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F \varrho_F^{\mathrm{I}}[\![h]\!][\![v]\!]\,\mathrm{d}F \\ \quad + \displaystyle\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F \varrho_F^{\mathrm{D}}hv\,\mathrm{d}F + \sum_{F\in\mathcal{F}_h^{S}}\int_F \varrho_F^{\mathrm{D}}\mathbb{1}_{\mathrm{S}}(h)hv\,\mathrm{d}F = \sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \varrho_F^{\mathrm{D}}h_{\mathrm{D}}v\,\mathrm{d}F \\ \quad + \displaystyle\sum_{F\in\mathcal{F}_h^{S}}\int_F \varrho_F^{\mathrm{D}}\mathbb{1}_{\mathrm{S}}(h)zv\,\mathrm{d}F, \\ \displaystyle\sum_{E\in\mathcal{E}_h}\int_E hv\,\mathrm{d}E = \sum_{E\in\mathcal{E}_h}\int_E h_0v\,\mathrm{d}E. \end{cases} \tag{II.25}$$

II.2.4 Symmetrization

The formulation is not symmetric between the unknown h and the test function v. Making the formulation symmetrical can be addressed in order to preserve the natural symmetry in the discrete diffusion operator, to use appropriate solvers or to enhance numerical properties of the formulation. This can be done by adding the corresponding symmetric term, $-\displaystyle\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F \{\!\!\{\mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}\}\!\!\}[\![h]\!]\,\mathrm{d}F$, which vanishes because $[\![h]\!]=0$ if h is a sufficiently smooth function. The terms:

$$-\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}h\,\mathrm{d}F - \sum_{F\in\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}h\,\mathrm{d}F, \tag{II.26}$$

are also added to get symmetry with their counterpart:

$$-\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}h_{\mathrm{D}}\,\mathrm{d}F - \sum_{F\in\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}z\,\mathrm{d}F. \tag{II.27}$$

Finally, the space semidiscretization reads:

$$\text{find } h\in S_p(\mathcal{E}_h^n) \text{ such that } \forall v\in S_p(\mathcal{E}_h^n),\quad m_{h,n}(\partial_t\theta(h-z),v)+a_{h,n}(h,v)=l_{h,n}(v), \tag{II.28}$$

where the DG bilinear forms, m and a, and linear form l are:

$$m_{h,n}(\partial_t\theta(h-z),v) = \sum_{E\in\mathcal{E}_h}\int_E \partial_t\theta(h-z)v\,\mathrm{d}E, \tag{II.29}$$

$$\begin{aligned}
a_{h,n}(h,v) = {} & \sum_{E\in\mathcal{E}_h}\int_E \mathbb{K}(h-z)\nabla h\cdot\nabla v\,\mathrm{d}E - \sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F \{\!\{\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}\}\!\}[\![v]\!]\,\mathrm{d}F \\
& + \sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F \varrho_F^{\mathrm{I}}[\![h]\!][\![v]\!]\,\mathrm{d}F - \Theta\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F \{\!\{\mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}\}\!\}[\![h]\!]\,\mathrm{d}F \\
& - \sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F + \sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \varrho_F^{\mathrm{D}}hv\,\mathrm{d}F \\
& - \Theta\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}h\,\mathrm{d}F \\
& - \sum_{F\in\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F + \sum_{F\in\mathcal{F}_h^{S}}\int_F \varrho_F^{\mathrm{D}}\mathbb{1}_{\mathrm{S}}(h)hv\,\mathrm{d}F \\
& - \Theta\sum_{F\in\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}h\,\mathrm{d}F,
\end{aligned} \tag{II.30}$$

$$\begin{aligned}
l_{h,n}(v) = {} & \sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \varrho_F^{\mathrm{D}}h_{\mathrm{D}}v\,\mathrm{d}F - \Theta\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F \mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}h_{\mathrm{D}}\,\mathrm{d}F - \sum_{F\in\mathcal{F}_h^{\mathrm{N}}}\int_F q_{\mathrm{N}}v\,\mathrm{d}F \\
& + \sum_{F\in\mathcal{F}_h^{S}}\int_F \varrho_F^{\mathrm{D}}\mathbb{1}_{\mathrm{S}}(h)zv\,\mathrm{d}F - \Theta\sum_{F\in\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}z\,\mathrm{d}F.
\end{aligned} \tag{II.31}$$

$\Theta = \{-1, 0, 1\}$ is a constant leading to the different symmetric versions of the DG formulation. In Eq. (II.29), Eq. (II.29) and Eq. (II.29), the terms are coloured differently according to their nature. Blue colour corresponds to element volume terms and green colour to interior face terms. Other colours are for boundary face terms: magenta is for Neumann boundary condition, red for Dirichlet boundary condition and orange for seepage boundary condition.

		Symmetrization		
		$\Theta = -1$	$\Theta = 0$	$\Theta = 1$
Penalization	$\forall F \in \mathcal{F}_h, \sigma_F^{\mathrm{I}} = \sigma_F^{\mathrm{D}} = 0$	OBB method	-	global element method
	$\forall F \in \mathcal{F}_h, \sigma_F^{\mathrm{I}} \neq 0, \sigma_F^{\mathrm{D}} \neq 0$	NIPG	IIPG	SIPG

NIPG: Non-symmetric Interior Penalty Galerkin SIPG: Symmetric Interior Penalty Galerkin
IIPG: Incomplete Interior Penalty Galerkin OBB method: Oden-Baumann-Babuška method

Table II.1 – Different types of discontinuous Galerkin methods.

Several DG methods can be derived according to the penalty weights and the coefficient of symmetry as listed in Tab. II.1. Their numerical properties are slightly different. Some basic properties are listed in Tab. II.2. Further information is available in [26, 27]. The IIPG method will be mainly chosen throughout the thesis because there is no additional terms from symmetrization which keeps the scheme simple which is interesting when the diffusion operator is nonlinear [27].

	Coercivity (stability)	Error estimation in L^2-norm	Remarks
IIPG	$\sigma > \sigma_*$	Suboptimal	Simplest version
SIPG	$\sigma > \sigma_*$	Optimal	Symmetric consistent version
NIPG	$\forall \sigma$	Suboptimal	Optimal error estimation with superpenalization

Table II.2 – Selected properties of some of the discontinuous Galerkin methods. σ_* is a threshold value for the penalty parameters σ.

The seepage boundary condition is treated numerically thanks to the indicator function $\mathbb{1}_{\mathrm{S}}$ which alternates dynamically between the Dirichlet part when $\mathbb{1}_{\mathrm{S}} = 1$ and the Neumann part for $\mathbb{1}_{\mathrm{S}} = 0$, directly in the usual weak framework of the DG methods. This treatment can model any physical situations, like multiple seepage faces simultaneously, because it does not require any assumption about the mesh or seepage face [77].

This treatment was also done by Dolejší *et al.* in [108] where they pointed out this is a great benefit compared to conforming FE methods which need to modify the FE space or use numerical tricks, resulting in difficult implementation. However, Dolejší *et al.* smoothed the indicator function $\mathbb{1}_F$ to avoid a brutal switch between the values 0 and 1 [108]. In this work, no regularization is performed because numerical experiments do not show major convergence solver problems. This makes the numerical scheme even more simple.

II.2.5 Algebraic representation

At a discrete level, the broken Sobolev space is approximated by a finite-dimensional subspace taken to be $S_p(\mathcal{E}_h^n)$ which is defined in Eq. (II.9). The number of global degrees of freedom in the mesh $\mathcal{E}_h$ is:

$$N_{\mathrm{dof}} := \dim S_p(\mathcal{E}_h^n) = \sum_{E \in \mathcal{E}_h^n} N_{\mathrm{dof}}^E, \tag{II.32}$$

where N_{dof}^E denotes the number of local degrees of freedom, that is to say those belonging to the element E. N_{dof}^E can vary from one element to another. For instance, it may depend on the chosen polynomial basis, the local polynomial degree p_E and even on the element shape (triangle, quadrangle, *etc*). The sought solution which could have be written as h_h to highlight its discrete nature. It is expanded in the global polynomial basis functions Φ of the space $S_p(\mathcal{E}_h^n)$ constructed from local basis functions ϕ of $S_p(E)$:

$$\forall t \in (0,T), \forall \boldsymbol{x} \in \Omega, \quad h(\boldsymbol{x},t) = \sum_{k=1}^{N_E} \sum_{l=1}^{N_{\mathrm{dof}}^E} \chi_k^l(t) \Phi_k^l(\boldsymbol{x}) \text{ where } \Phi_k^l(\boldsymbol{x}) = \begin{cases} \phi_k^l(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in E \\ 0 & \text{otherwise} \end{cases}. \tag{II.33}$$

The degrees of freedom are denoted by χ_k^l, and are functions of time. The algebraic representation is interesting to make the link with programming described in Section II.7. It helps writing the matrix system originated from Eq. (II.28) and understanding the choice of polynomial basis functions.

II.3 Time discretization

To get the fully discrete DG formulation, the time derivative has now to be discretized. Such a procedure which takes care of space discretization before time discretization is called the method of lines.

It is used commonly by others studies for Richards' equation, see *e.g.* [20, 21, 23]. Richards' equation is a stiff equation because of constitutive laws which usually dismisses explicit methods for time integration even though this choice can be done, for example in [124] under justifications of better parallel computing performances. For stability reasons, implicit methods are used for Richards' equation : backward Euler method, Crank-Nicolson method, backward differentiation formula methods, diagonally implicit Runge-Kutta (DIRK), *etc*. The use of low-order schemes is well-established but a trend can be oserved for higher order time schemes [105].

The following notation will be used for any function $u \in L^2(0,T;\Omega)$:

$$\forall n \in \mathbb{N}, 0 \leq n < N, u^n(\boldsymbol{x}) \coloneqq u(\boldsymbol{x}, t_n). \tag{II.34}$$

II.3.1 ϑ-Method

The ϑ-method (theta-method) is used to integrate partial differential equations in time and depends on the parameter $\vartheta \in \mathbb{R}$ such that $0 \leq \vartheta \leq 1$. It is a generalization of the Euler methods since it represents the convex linear combination of the explicit Euler method (forward Euler method) and the implicit Euler method (backward Euler method). The later are simple and popular method of order 1. Besides, ϑ-method also encompasses the Crank-Nicolson method which can be viewed as the arithmetic mean of the Euler methods. Crank-Nicolson method is based on the trapezoidal rule providing second order convergence in time.

The DG formulation in Eq. (II.28) is integrated over T^{n+1} and the time integral is approximated by the ϑ-method:

find a sequence of $(h^n)_{0\leq n<N} \in S_p(\mathcal{E}_h^n)$ such that

$$\begin{cases} h^0 = h_0, \\ \forall v \in S_p(\mathcal{E}_h^n), \quad m_{h,n}\left(\dfrac{\theta(h^{n+1}) - \theta(h^n)}{\tau^{n+1}}, v\right) & +(1-\vartheta)(a_{h,n}(h^n, v) - l_{h,n}(v, t^n)) \\ & +\vartheta\Big(a_{h,n}\Big(h^{n+1}, v\Big) - l_{h,n}\Big(v, t^{n+1}\Big)\Big) = 0. \end{cases} \tag{II.35}$$

It can be noticed that $\vartheta = 1$ leads to the implicit Euler method, $\vartheta = 0$ to the explicit Euler method and $\vartheta = \dfrac{1}{2}$ to the Crank-Nicolson method.

Forward Euler scheme must be excluded because it is conditionally stable (explicit) under time steps restriction known as the CFL condition unless using tiny time stepsizes which is not viable. Crank-Nicolson scheme should be used carefully: despite being of second-order, it provides less accurate results than backward Euler scheme of order one. This is due to its semi-implicit nature which may cause oscillations. The latter may persist in the solution because they are damping weakly. Authors notice oscillations occur if time stepsize is too large or if there are jump discontinuities in the initial solution or between initial condition and boundary conditions [125]. This situation is inherent to either Richards' equation or DG method, which may explain why Crank-Nicolson scheme is not a good choice.

Even though the backward Euler scheme has error of order one in time, it will be particularly used for Richards' equation in this study because the implicit nature of the scheme provides good stability properties. Indeed, the backward Euler scheme shows a huge region of absolute stability, making it A-stable, and even L-stable, which is suitable to integrate stiff equations. That is why, together with

its simple implementation, this scheme is widespread for Richards' equation where good convergence behaviour is observed in [17].

II.3.2 Backward difference formula methods

Backward differentiation formula methods (BDF) are implicit numerical integration methods especially used for stiff differential equations thanks to their wide region of stability. These linear multistep methods use q already computed solutions to produce a method of order q. BDF methods with $q > 6$ are not stable. Analysis can be found in [126, 127]. The 1-step BDF method is the backward Euler scheme. For this study, BDF methods are interesting because they provide high-order accuracy in time which is suitable if one wants to reach high-order in space without loss of gain. In other words, the BDF methods will allow to balance space and time errors. Besides, coefficients of BDF methods are easily computed and the cost of integration stays almost constant when order gets higher since the only requirements is to have access to older already computed solutions, making them competitive compared to implicit Runge-Kutta methods.

The DG formulation (II.28) is integrated over T^{n+1} and the time integral is approximated by the BDF methods of $1 \leq q \leq 6$ steps with variable time steps [128]. The BDF method of order q is obtained from the Newton interpolation polynomial of degree q interpolating h^j at time t^j for $j = n+1, \dots, n+1-q$. Divided differences are useful to express compactly variable stepsize BDF methods but also to compute efficiently and automatically coefficients using recursion. The backward divided difference of a given function y is defined by a recursive division process defined as:

$$\begin{cases} \left[y^{n+1}\right] := y^{n+1}, \\ \left[y^{n+1}, \dots, y^{n+1-j}\right] := \dfrac{\left[y^{n+1}, \dots, y^{n+2-j}\right] - \left[y^{n}, \dots, y^{n+1-j}\right]}{t^{n+1} - t^{n+1-j}}. \end{cases} \tag{II.36}$$

The implicit BDF method corresponds to:

$$\partial_t \theta(h - z) = \sum_{j=1}^{q} \left(\prod_{k=1}^{j-1} \left(t^{n+1} - t^{n+1-k}\right) \right) \left[\theta\left(h^{n+1} - z\right), \dots, \theta\left(h^{n+1-j} - z\right)\right] \tag{II.37}$$

$$= \sum_{k=0}^{q} \frac{\alpha_{s,k}}{\tau^{n+1}} \theta\left(h^{n+1-k} - z\right), \tag{II.38}$$

where $\alpha_{s,k}$ are the coefficients of the linear combination obtained from the coefficients of the divided differences of h, where $\sum_{k=0}^{q} \alpha_{s,k} = 0$. This provides an approximation of $\partial_t \theta\left(h^{n+1} - z\right)$ based on the $\theta\left(h^{j} - z\right)$ and guarantees $\partial_t \theta\left(h^{n+1} - z\right) = 0$ if the problem is steady. Besides, the last coefficient is computed as the opposite sum of the other coefficients:

$$\alpha_{s,q} = -\sum_{k=0}^{q-1} \alpha_{s,k}. \tag{II.39}$$

This ensures that a constant function is exactly integrated at machine precision adjusting the coefficients with the largest round-off error since it involves the highest number of levels of divided differences. When the time stepsize is fixed, the coefficients $\alpha_{s,k}$ takes the classic values of the standard BDF method. They are given in Tab. II.3 because they may serve as a test for computation.

$\boldsymbol{q}$								
	1	1	-1					
	2	$3/2$	-2	$1/2$				
	3	$11/6$	-3	$3/2$	$-1/3$			
	4	$25/12$	-4	3	$-4/3$	$1/4$		
	5	$137/60$	-5	5	$-10/3$	$5/4$	$-1/5$	
	6	$49/20$	-6	$15/2$	$-20/3$	$15/4$	$-6/5$	$1/6$
$\boldsymbol{\alpha_{s,k}}$		**0**	**1**	**2**	**3**	**4**	**5**	**6**
		$\boldsymbol{k}$						

Table II.3 – Coefficients $\alpha_{q,k}$ for all BDF methods with fixed time stepsize.

Then, the complete discretized system with BDF methods reads:

find a sequence of $(h^n)_{0\leq n<N} \in S_p(\mathcal{E}_h^n)$ such that

$$\begin{cases} h^0 = h_0, \\ \forall v \in S_p(\mathcal{E}_h^n), \quad m_{h,n}\left(\sum_{k=0}^{q} \frac{\alpha_{q,k}}{\tau^{n+1}}\theta\left(h^{n+1-k} - z\right), v\right) + a_{h,n}\left(h^{n+1}, v\right) = l_{h,n}\left(v, t^{n+1}\right). \end{cases} \tag{II.40}$$

However, stability of BDF methods has some restrictions. Only BDF methods of order 1 and 2 are A-stable according to the second Dahlquist barrier for implicit linear multistep methods as defined and proved in [129]. BDF methods of order 3 to 6 are conditionally stable or A(α)-stable where α decreases with the order. BDF methods of order $q > 6$ are unconditionally unstable so they cannot be used. To strengthen stability, one can use variable time steps. Roughly speaking, reducing time step length gives stability while increasing time step length degrades stability properties. In practice, time step size variation are bounded above by a value called the swing factor to guarantee stability and give robustness. The restriction becomes more severe as the order q gets higher. Strategies are described in [128, 130] to adapt the time stepsize but also to select the appropriate order of BDF method according to stability indicator. Some theoretical bounds for swing factor exist [131] but they are very severe since all sequences of time steps lengths are considered, included pathological cases. In practice, time stepping is usually constructed to vary smoothly so, in this study, following Hay *et al*. [128], practical maximum swing between two consecutive time steps are taken from Skelboe [132] and reported in Tab. II.4. They are still relatively pessimistic but favours stability with time stepping which can be adapted conveniently. The 1-BDF method is unconditionally stable because it is L-stable and has no swing limit. The 6-BDF method cannot increase rapidly time step lengths and thus is not very useful if time step has to adapt constantly during the simulation. Despite these stability restrictions, BDF methods behave better for stiff problems than other methods like Adams methods because they show a stiff decay property. This means BDF methods damp quickly numerical errors due to oscillatory components in the solution if they are applied in their stability region. This is sometimes referred as neutral stability and is the reason why BDF methods are promoted for stiff problems.

Order q	1	2	3	4	5	6
Maximum swing τ^{n+2}/τ^{n+1}	-	2.6	1.9	1.5	1.2	1.05

Table II.4 – Maximum swing τ^{n+2}/τ^{n+1} for BDF methods with variable time stepsize.

BDF methods are very interesting for Richards' equation known to be stiff in some cases. For instance, Crank-Nicolson scheme is the ϑ-scheme for $\vartheta = \frac{1}{2}$ and corresponds to Adams-Moulton method of order 2. Even though the method is semi-implicit, spurious oscillations may occur for large time stepsize under certain conditions because oscillations are not damped fast enough for all decaying components.

II.3.3 Initialization

Initialization takes care of prescribing initial condition to start the first time step. This corresponds of the first line of Eq. (II.35) or Eq. (II.40). A direct and simple way is to write the corresponding discontinuous weak formulation:

$$\forall v \in S_p\big(\mathcal{E}_h^0\big), \sum_{E\in\mathcal{E}_h}\int_E h^0 v \,\mathrm{d}E = \sum_{E\in\mathcal{E}_h}\int_E h_0 v \,\mathrm{d}E. \tag{II.41}$$

This consists in solving a system enforced by the right-hand side, which can be formally written as:

$$\text{find } h^0 \in S_p\big(\mathcal{E}_h^0\big) \text{ such that } \forall v \in S_p\big(\mathcal{E}_h^0\big), \quad m_{h,0}\big(h^0, v\big) = f_{h,0}(v), \tag{II.42}$$

where m is the bilinear form already defined and f is the following linear form:

$$f_{h,0}(v) = \sum_{E\in\mathcal{E}_h}\int_E h_0 v \,\mathrm{d}E. \tag{II.43}$$

One limiting aspect of BDF methods is for initialization seeing that a s-step BDF method needs s previous solutions to be computed. Then, the first time step given by the user is divided in s sub-steps. Starting with 1-BDF method, the order is progressively increased at each sub-steps to reach the specified s-steps BDF method for the user-defined initial time steps.

If one works with BDF methods not based on variable stepsize, initialization and a change of time stepsize has to be treated with a more complicated technique. For a scheme with q-step BDF method, the first of the next time steps is subdivided into as many times as necessary to compute the sub-time steps with an increasing sequence of lower-order BDF methods. The remaining next macro-time steps are still subdivided while the number of computed solution, stored gradually, is not sufficient to go on with the q-step BDF method.

II.4 Local mass balance

Mass conservation is an important – necessary – property for flow problems in porous media since it can cause poor quality results or instabilities [18]. A numerical solution of Richards' equation is a good approximation if it is mass conservative [85, 133]. However, the inverse is not true [85] and additional criterion should assess whether a solution is correct or not such as monotony and maximum principle [91].

The mixed formulation of Richards' equation is usually preferred because it is *physically* mass conservative contrary to the other formulations [85] due to the chain differentiation rule for the time derivative. Then, discretization scheme should *numerically* preserve mass balance, at least globally, and if possible, locally. Unlike FE methods which are globally mass conservative over the computational domain, DG methods are particularly suitable for porous media flow problems [26] because they satisfy naturally local mass conservation, that is to say for each element.

In order to introduce this property, let fix an element E and a test function $v \in \mathcal{D}_p(\mathcal{E}_h)$ that is equal to one on E and vanishes outside of E. This is possible with DG methods due to the lack of continuity constraints between the elements. The equation (II.28) reduces to:

$$\begin{aligned}
&\int_E \partial_t \theta(h-z)\,\mathrm{d}E - \sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{I}}}\int_F \{\!\!\{\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}\}\!\!\}\,\mathrm{d}F - \sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{D}}}\int_F \mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}\,\mathrm{d}F\\
&- \sum_{F\in\partial E\cap\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n_F}\,\mathrm{d}F + \sum_{F\partial E\cap\in\mathcal{F}_h^{\mathrm{I}}}\int_F \varrho_F^{\mathrm{I}}[\![h]\!]\,\mathrm{d}F + \sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{D}}}\int_F \varrho_F^{\mathrm{D}} h\,\mathrm{d}F\\
&+ \sum_{F\in\partial E\cap\mathcal{F}_h^{S}}\int_F \varrho_F^{\mathrm{D}}\mathbb{1}_{\mathrm{S}}(h)h\,\mathrm{d}F = \sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{D}}}\int_F \varrho_F^{\mathrm{D}} h_{\mathrm{D}}\,\mathrm{d}F + \sum_{F\in\partial E\cap\mathcal{F}_h^{S}}\int_F \varrho_F^{\mathrm{D}}\mathbb{1}_{\mathrm{S}}(h)z\,\mathrm{d}F\\
&- \sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{N}}}\int_F q_{\mathrm{N}}\,\mathrm{d}F.
\end{aligned}$$

Then, let integrate the previous equation over time T^{n+1} with $n \in \mathbb{N}, 0 \leq n < N$. Here it is done with BDF methods to provide high-order accuracy but the procedure is similar with ϑ-method. Reorganizing the terms, it yields:

$$m_{\mathrm{add}}\left(t^{n+1}\right) = m_{\mathrm{flux}}\left(t^{n+1}\right) + m_{\mathrm{art}}\left(t^{n+1}\right), \tag{II.44}$$

where the quantities m_{add}, m_{flux} and m_{art} are defined by:

$$m_{\mathrm{add}}\left(t^{n+1}\right) = \int_E \left(\sum_{k=0}^{q}\alpha_{q,k}\theta\left(h^{n+1-k}-z\right)\right)\mathrm{d}E, \tag{II.45}$$

$$\begin{aligned}
m_{\mathrm{flux}}\left(t^{n+1}\right) = \sum_{l=0}^{n}\tau^{l+1}\Bigg(&\sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{I}}}\int_F\left\{\!\!\left\{\mathbb{K}\left(h^l-z\right)\nabla h^l\cdot\boldsymbol{n_F}\right\}\!\!\right\}\mathrm{d}F\\
&+\sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{D}}}\int_F \mathbb{K}\left(h^l-z\right)\nabla h^l\cdot\boldsymbol{n_F}\,\mathrm{d}F - \sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{N}}}\int_F q_{\mathrm{N}}\,\mathrm{d}F\\
&+\sum_{F\in\partial E\cap\mathcal{F}_h^{S}}\int_F \mathbb{1}_{\mathrm{S}}\left(h^l\right)\mathbb{K}\left(h^l-z\right)\nabla h^l\cdot\boldsymbol{n_F}\,\mathrm{d}F\Bigg),
\end{aligned} \tag{II.46}$$

$$\begin{aligned}
m_{\mathrm{art}}\left(t^{n+1}\right) = \sum_{l=0}^{n}\tau^{l+1}\Bigg(&-\sum_{F\partial E\cap\in\mathcal{F}_h^{\mathrm{I}}}\int_F \varrho_F^{\mathrm{I}}[\![h^l]\!]\,\mathrm{d}F - \sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{D}}}\int_F \varrho_F^{\mathrm{D}} h^l\,\mathrm{d}F\\
&+\sum_{F\in\partial E\cap\mathcal{F}_h^{\mathrm{D}}}\int_F \varrho_F^{\mathrm{D}} h_{\mathrm{D}}\,\mathrm{d}F - \sum_{F\in\partial E\cap\mathcal{F}_h^{S}}\int_F \varrho_F^{\mathrm{D}}\mathbb{1}_{\mathrm{S}}\left(h^l\right)h^l\,\mathrm{d}F + \sum_{F\in\partial E\cap\mathcal{F}_h^{S}}\int_F \varrho_F^{\mathrm{D}}\mathbb{1}_{\mathrm{S}}\left(h^l\right)z\,\mathrm{d}F\Bigg).
\end{aligned} \tag{II.47}$$

The total additional mass in the domain m_{add} is a measure with respect to the initial mass in the system. If one wants to incorporate a sink/source term, it should be added to this quantity to take into account mass destruction/creation. The total net flux into the domain m_{flux} is a measure of the mass flow rate passing through the boundaries ∂E. The additional artificial mass in the domain m_{art} is a measure of the pure numerical mass created by penalty parameters. This quantity is vanishing if the penalty is zero. It can be removed if needed since this artificial mass can be exactly computed. For Richards' equation with primal DG methods, mass balance β_{m} is defined with the following ratio:

$$\beta_{\mathrm{m}}\left(t^{n+1}\right) := \frac{m_{\mathrm{add}}(t^{n+1})}{m_{\mathrm{flux}}(t^{n+1}) + m_{\mathrm{art}}(t^{n+1})}. \tag{II.48}$$

Error on mass balance ϵ_{m} can be calculated and it is sometimes expressed as a rate τ_{m}:

$$\epsilon_{\mathrm{m}}\left(t^{n+1}\right) := \left|1 - \beta_{\mathrm{m}}\left(t^{n+1}\right)\right|, \tag{II.49}$$

$$\tau_{\mathrm{m}}\left(t^{n+1}\right) := 100\epsilon_{\mathrm{m}}\left(t^{n+1}\right). \tag{II.50}$$

Mass balance must be equal to one. It means that the change of the fluid mass in the domain between two time steps should be equal to the integral of the mass flux over the domain boundary. Disregarding artificial mass, if the numerical approximation of the velocity - the total flux - is not locally conservative compared to water content change, the numerical solution of Richards' equation becomes unstable after a few time steps. This equilibrium could be violated because of inaccurate numerical integration for example, see [85]. Mass balance is assessed in Section IV.2 and Section IV.3.

II.5 Nonlinear iterative process

Richards' equation is a nonlinear equation usually solved by an iterative procedure such as fixed-point iteration (also called Picard's) or Newton-Raphson method whose choice is determining for computation time performances and convergence. Lots of derived methods exist: modified Newton-Raphson, mixed Picard–Newton, quasi-Newton, L-scheme, *etc*. Studies have been carried out to compare these methods, and others, for solving Richards' equation for various problems [17, 18, 30]. They emphasize that fixed-point iteration and even Newton-Raphson scheme are very sensitive and do not converge systematically according to Richards' equation's formulations, initial and boundary conditions and because of the nonlinearities of constitutive laws. In particular, the nonlinear iterative solver can oscillate between two solutions. However, this methods are problem-dependent and a conclusion cannot be generalized.

II.5.1 Newton-Raphson method

For this study, no source/sink term is considered but if one exists and depends on the solution, it should be incorporated in the linearization process and, then the linear form l may depend on h. A damped Newton-Raphson method and a fixed-point method were implemented. Let m be the number of the nonlinear iteration. The residual of Eq. (II.40) is expressed as:

$$r_{h,n}\left(h^{n+1}, v\right) := m_{h,n}\left(\sum_{k=0}^{q} \frac{\alpha_{s,k}}{\tau^{n+1}}\theta\left(h^{n+1-k} - z\right), v\right) + a_{h,n}\left(h^{n+1}, v\right) - l_{h,n}\left(v; t^{n+1}\right). \tag{II.51}$$

Linearization

The difference between two successive iterations is written $\delta_h^{n+1,m} := h^{n+1,m+1} - h^{n+1,m}$ so the iterative procedure reads:

$$\begin{cases} \dfrac{\mathrm{d}r_{h,n}(h^{n+1,m}, v)}{\mathrm{d}h^{n+1,m}}\delta_h^{n+1,m} = -r_{h,n}\left(h^{n+1,m}, v\right), \\ h^{n+1,m+1} = h^{n+1,m} + \delta_h^{n+1,m}. \end{cases} \tag{II.52}$$

In Eq. (II.52), $\dfrac{\mathrm{d}r_{h,n}(h^{n+1,m}, v)}{\mathrm{d}h^{n+1,m}}$ corresponds to the Jacobian matrix, also known to be the tangent stiffness matrix. This term can be calculated with asymptotic developments. Indeed, provided that

$r_{h,n}\Big(h^{n+1,m+1},v\Big)=0$, one can write:

$$\begin{aligned} r_{h,n}\Big(h^{n+1,m+1},v\Big) &= r_{h,n}\Big(h^{n+1,m}+\delta_h^{n+1,m},v\Big) \\ &= r_{h,n}\Big(h^{n+1,m},v\Big)+\frac{\mathrm{d}r_{h,n}(h^{n+1,m},v)}{\mathrm{d}h^{n+1,m}}\delta_h^{n+1,m}+o\Big(\delta_h^{n+1,m}\Big) \qquad \text{(II.53)} \\ \iff \frac{\mathrm{d}r_{h,n}(h^{n+1,m},v)}{\mathrm{d}h^{n+1,m}}\delta_h^{n+1,m} &= -r_{h,n}\Big(h^{n+1,m},v\Big)+o\Big(\delta_h^{n+1,m}\Big). \qquad \text{(II.54)} \end{aligned}$$

Then, the expression of Jacobian matrix of Eq. (II.52) can be deduced by identifying the term-wise asymptotic developments of $r_{h,n}\Big(h^{n+1,m+1},v\Big)$ with Eq. (II.53). To do this, the following asymptotic developments should be used:

$$\theta\Big(h^{n+1,m}+\delta_h^{n+1,m}-z\Big)=\theta\Big(h^{n+1,m}-z\Big)+\frac{\mathrm{d}\theta(h^{n+1,m}-z)}{\mathrm{d}h^{n+1,m}}\delta_h^{n+1,m}+o\Big(\delta_h^{n+1,m}\Big), \qquad \text{(II.55)}$$

$$\mathbb{K}\Big(h^{n+1,m}+\delta_h^{n+1,m}\Big)=\mathbb{K}\Big(h^{n+1,m}\Big)+\frac{\mathrm{d}\mathbb{K}(h^{n+1,m})}{\mathrm{d}h^{n+1,m}}\delta_h^{n+1,m}+o\Big(\delta_h^{n+1,m}\Big). \qquad \text{(II.56)}$$

Identification gives the following expression for the Jacobian matrix:

$$\begin{aligned} \frac{\mathrm{d}r_{h,n}(h^{n+1,m},v)}{\mathrm{d}h^{n+1,m}}\delta_h^{n+1,m} &= \sum_{E\in\mathcal{E}_h}\int_E\frac{\mathrm{d}\theta(h^{n+1,m}-z)}{\mathrm{d}h}\delta_h v\,\mathrm{d}E+\sum_{E\in\mathcal{E}_h}\int_E\mathbb{K}(h-z)\nabla\delta_h\cdot\nabla v\,\mathrm{d}E \qquad \text{(II.57)} \\ &+\sum_{E\in\mathcal{E}_h}\int_E\delta_h\frac{\mathrm{d}\mathbb{K}(h^{n+1,m}-z)}{\mathrm{d}h}\nabla h\cdot\nabla v\,\mathrm{d}E-\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F\{\!\!\{\mathbb{K}(h-z)\nabla\delta_h\cdot\boldsymbol{n_F}\}\!\!\}[\![v]\!]\,\mathrm{d}F \\ &-\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F\left\{\!\!\left\{\delta_h\frac{\mathrm{d}\mathbb{K}(h^{n+1,m}-z)}{\mathrm{d}h}\nabla h\cdot\boldsymbol{n_F}\right\}\!\!\right\}[\![v]\!]\,\mathrm{d}F-\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F\mathbb{K}(h-z)\nabla\delta_h\cdot\boldsymbol{n_F}v\,\mathrm{d}F \\ &-\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F\delta_h\frac{\mathrm{d}\mathbb{K}(h^{n+1,m}-z)}{\mathrm{d}h}\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F-\sum_{F\in\mathcal{F}_h^{S}}\int_F\mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla\delta_h\cdot\boldsymbol{n_F}v\,\mathrm{d}F \\ &-\sum_{F\in\mathcal{F}_h^{S}}\int_F\mathbb{1}_{\mathrm{S}}(h)\delta_h\frac{\mathrm{d}\mathbb{K}(h^{n+1,m}-z)}{\mathrm{d}h}\nabla h\cdot\boldsymbol{n_F}v\,\mathrm{d}F+\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F\varrho_F^{\mathrm{I}}[\![\delta_h]\!][\![v]\!]\,\mathrm{d}F \\ &+\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F\varrho_F^{\mathrm{D}}\delta_h v\,\mathrm{d}F+\sum_{F\in\mathcal{F}_h^{S}}\int_F\varrho_F^{\mathrm{D}}\mathbb{1}_{\mathrm{S}}(h)\delta_h v\,\mathrm{d}F-\Theta\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F\{\!\!\{\mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}\}\!\!\}[\![\delta_h]\!]\,\mathrm{d}F \\ &-\Theta\sum_{F\in\mathcal{F}_h^{\mathrm{I}}}\int_F\left\{\!\!\left\{\delta_h\frac{\mathrm{d}\mathbb{K}(h^{n+1,m}-z)}{\mathrm{d}h}\nabla v\cdot\boldsymbol{n_F}\right\}\!\!\right\}[\![h]\!]\,\mathrm{d}F-\Theta\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F\mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}\delta_h\,\mathrm{d}F \\ &-\Theta\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F\delta_h\frac{\mathrm{d}\mathbb{K}(h^{n+1,m}-z)}{\mathrm{d}h}\nabla v\cdot\boldsymbol{n_F}h\,\mathrm{d}F-\Theta\sum_{F\in\mathcal{F}_h^{S}}\int_F\mathbb{1}_{\mathrm{S}}(h)\mathbb{K}(h-z)\nabla v\cdot\boldsymbol{n_F}\delta_h\,\mathrm{d}F \\ &-\Theta\sum_{F\in\mathcal{F}_h^{S}}\int_F\mathbb{1}_{\mathrm{S}}(h)\delta_h\frac{\mathrm{d}\mathbb{K}(h^{n+1,m}-z)}{\mathrm{d}h}\nabla v\cdot\boldsymbol{n_F}h\,\mathrm{d}F \\ &+\Theta\sum_{F\in\mathcal{F}_h^{\mathrm{D}}}\int_F\delta_h\frac{\mathrm{d}\mathbb{K}(h^{n+1,m}-z)}{\mathrm{d}h}\nabla v\cdot\boldsymbol{n_F}h_{\mathrm{D}}\,\mathrm{d}F \\ &+\Theta\sum_{F\in\mathcal{F}_h^{S}}\int_F\mathbb{1}_{\mathrm{S}}(h)\delta_h\frac{\mathrm{d}\mathbb{K}(h^{n+1,m}-z)}{\mathrm{d}h}\nabla v\cdot\boldsymbol{n_F}z\,\mathrm{d}F. \qquad \text{(II.58)} \end{aligned}$$

In Eq. (II.58), the terms are coloured differently. Blue colour terms are zero-order since they does not involve any derivative. They match exactly the bilinear form a. Whereas red colour terms are one-order because they hold derivatives. Such a distinction is useful to know the definition of the tangent stiffness matrix in case of the fixed-point method described in Section II.5.2.

Derivatives of constitutive laws (hydraulic properties) present in the Jacobian matrix can be computed analytically. Nevertheless, the resulting expressions are heavy. That is why derivatives are approximated using the central difference formula for a function u evaluated at the point h:

$$\frac{\mathrm{d}u(h)}{\mathrm{d}h} = \frac{u(h+\varepsilon) - u(h-\varepsilon)}{2\varepsilon}, \tag{II.59}$$

where $\varepsilon = u(h)\sqrt{\epsilon_{\mathrm{pr}}}$ with ϵ_{pr} the machine epsilon if $u(h) \neq 0$. Otherwise, $\varepsilon = \sqrt[3]{\epsilon_{\mathrm{pr}}}$. This rule of thumb aims for a small value of ε but not too small due to roundoff errors in the representation of real numbers.

Stopping criteria

Both Newton-Raphson and fixed-point methods require a first guess for the iterative process which is taken from the solution of the last time step. One important choice, determining accuracy for nonlinear iterative process, is the stopping criterion. For Richards' equation, the stopping criterion can be specified in terms of absolute error for pressure head or water content between two successive iterations [134]. For this study, the stopping criteria read:

$$\frac{\|r_{h,n}(h^{n+1,m}, v)\|_{L^2(\Omega)}}{\|a_{h,n}(h^{n+1,m}, v)\|_{L^2(\Omega)}} < \varepsilon_1 \quad \text{and} \quad \frac{\left\|\delta_h^{n+1,m}\right\|_{L^2(\Omega)}}{\|h^{n+1,m}\|_{L^2(\Omega)}} < \varepsilon_2, \tag{II.60}$$

where $\varepsilon_1, \varepsilon_2$ are a user-defined tolerances. These two criteria are relative in order to be independent from the characteristic quantities of the problem. The first criterion is residual and the second one is incremental.

II.5.2 Approximation techniques

Relaxation and damping

The convergence of the nonlinear process is known to be difficult for Richards' equation. In particular, the nonlinear iterative solver can oscillate between two solutions. This is due to steep nonlinearities of hydraulic properties in the capillary region under some circumstances. Many techniques try to avoid this situation to reach robustness. Most of them involve adaptive time stepping or under-relaxation. Adaptive time techniques are discussed in the next section. Under-relaxation techniques can affect:

- the nonlinear iterative solution with damping or Anderson acceleration;
- the hydraulic conductivity with chord slope approximation, polynomial spline or regularization;
- the nonlinear method with tangent matrix approximation or clever initial guessing.

General studies [17, 18, 30] show that none of these techniques can solely make the solver robust. Some tests tend to indicate that a Newton-Raphson method in combination with a fixed-point method for the first iterations is sufficiently robust and efficient [18].

The damped Newton-Raphson method relaxes the increment $\delta_h^{n+1,m}$ in Eq. (II.52) while the new residual is greater than the previous one. The relaxation of the increment is done thanks to line search. Such a procedure is described in [108] and has been implemented.

Fixed-point method

Newton-Raphson methods converge quadratically but they often fail for parabolic degenerate problem where the Jacobian matrix might become singular and because of poor initial guess (locally convergent) [17, 18, 30]. Besides, Newton-Raphson methods require more computational resources. Therefore, fixed-point iteration methods, also known as Picard's method, are robust and popular techniques for Richards' equation even if they converge linearly. This is demonstrated in Chapter IV, particularly on the case of Section IV.7.

Underlining ideas are already described in [18] and have just been adapted here to the DG system. Roughly speaking, a fixed-point method is approximated from the Newton-Raphson method by avoiding the first derivative terms inside the Jacobian matrix, that is to say only the terms in blue are kept in Eq. (II.58).

II.6 Adaptive time stepping

Time adaptation is motivated by the convergence of the nonlinear solver. Indeed, nonlinearities may be so strong that the iterative process may oscillate between two solution states. This is even more the case because of stiff initialization, poor initial guess or severe boundary conditions. On one hand, transient simulations have difficulties to converge if the time step is too large but, on the other hand, shorter time steps mean more time steps and so, a longer computational time. That is the reason why time adaptation is very attractive and common for Richards' equation. Different strategies can be used to adjust the time step [23]. They are either heuristic, mainly based on convergence performance of the nonlinear solver [19, 37, 135, 136], or rational and based on error control [20, 98, 137–139]. The latter ones are generally more efficient but heuristic methods may outperform them in regards of their simple implementation and so they remains a relevant approach for Richards' equation.

For this study, the time step is adjusted heuristically according to the previous number of iterations N_{it} from the nonlinear solver such as [19, 135]. Let $n \in \mathbb{N}, 0 \leq n < N$. The simulations begins with a time step τ^1. The next time step τ^{n+2} is calculated from the previous one τ^{n+1} according to the time-stepping scheme:

$$\begin{cases} \tau^{n+2} = \begin{cases} \lambda_{\text{amp}}\tau^{n+1} & \text{if } N_{\text{it}} \leq m_{\text{it}}, \\ \tau^{n+1} & \text{if } m_{\text{it}} < N_{\text{it}} \leq M_{\text{it}}, \\ \lambda_{\text{red}}\tau^{n+1} & \text{if } M_{\text{it}} < N_{\text{it}} \leq W_{\text{it}}, \end{cases} \\ \tau^{n+1} = \lambda_{\text{red}}\tau^{n+1} \text{ if } W_{\text{it}} < N_{\text{it}} \text{ or if the solver has failed (time step is started again).} \end{cases} \tag{II.61}$$

The factors $0 < \lambda_{\text{red}} < 1$ and $1 < \lambda_{\text{amp}}$ as well as the threshold values $0 < m_{\text{it}} \leq M_{\text{it}} < W_{\text{it}}$ are prescribed by the user. A minimum time step is defined to avoid excessive small time step.

With this approach, the nonlinear solver is more robust because the time step is adjusted until success of convergence whatever τ^1. Thanks to the amplification/reduction coefficients, the time step is adjusted smoothly. Nevertheless, the method depends on fixed empirical parameters restricting the possible size of the new time step. This does not provide an optimal time step. The resulting loss/gain in computational time is difficult to assess in regards to the balance between the number of nonlinear iterations and time steps length.

II.7 Implementation choices

II.7.1 Overall programming

Rivage code

In order to solve Richards' equation and simulate flows in unsaturated porous media, the development of a program was undertaken from scratch to implement the numerical methods described in this thesis. The code is written in Fortran 2008 and compiled with *GFortran* version 10.2. It is called *Rivage* which stands for « RIchards pour des VAgues par méthodes de Galerkine discontinuEs » (Richards for waves by discontinuous Galerkin methods). The goal of *Rivage* code is to provide a generic tool based on the DG method. The structures of the program and the variables are intended to be reused for other model equations than Richards' equation. The code can already be employed for a large range of features: non-conforming, unstructured and hybrid mesh of triangles and parallelograms, three basis functions, quadrature rules exact up to degree 20, three solvers based on different matrix storage, a clever numbering, *etc*. Moreover, the code development tried to follow as much as possible the quality requirements for programming. Among others, the user has many possibilities to choose for the numerical methods, gathered in a parameter file to make it more usable. A number of errors will result in messages in the terminal for the user who can find additional comments in the code. Finally, the code is easily maintainable and modular which makes future extensions possible. The structure and the subroutines were especially designed to work in 3D, to add hyperbolic term with numerical flux and to be parallelized.

The execution work with different executable files which communicate between them with binary files called *rivbin* containing necessary data. Executable files consist in:

- *data2rivage* which converts parameter file *.prm* into binary data readable by *rivage*;
- *rivage* which is the main executable file. It performs calculations for the simulation;
- *rivage2visu* which converts binary data exported by *rivage* into data file under the format *.dat* readable by visualization softwares like Tecplot or VisIt;
- *adaptriv* which can modify the mesh from a binary file of *rivage*;
- *modifriv* which can modify the parameters of the simulation from a binary file of *rivage*.

In addition, one file of extension *.prm* is mandatory to execute the program. The format is specific to *Rivage* and it contains all the physical and numerical parameters of the problem to perform the simulation. Within this file, it is possible to create a mesh made up of block-based grids. If one wants to specify particular mesh or boundary conditions, then optional files may be considered. A file of extension *.msh* can be read to create an original mesh. This is at the format up to version 2.16.0 of *Gmsh*, an open source 3D FE mesh generator. In option, a file *.lim* can be provided to take into account dynamic Dirichlet boundary conditions monitored by data. They are written in a format only readable by *Rivage* and can serve as prescribing dynamic forcing boundary condition.

The core algorithm of *Rivage* is described in Fig. II.5. There are two embedded loops. The main one corresponds to the time cycle where each iteration goes through several steps: the solution of Richards' equation, adaptation of time stepsize, remeshing and eventually data export for visualization. The second loop concerns the nonlinear iterative process to get the solution of Richards' equation.

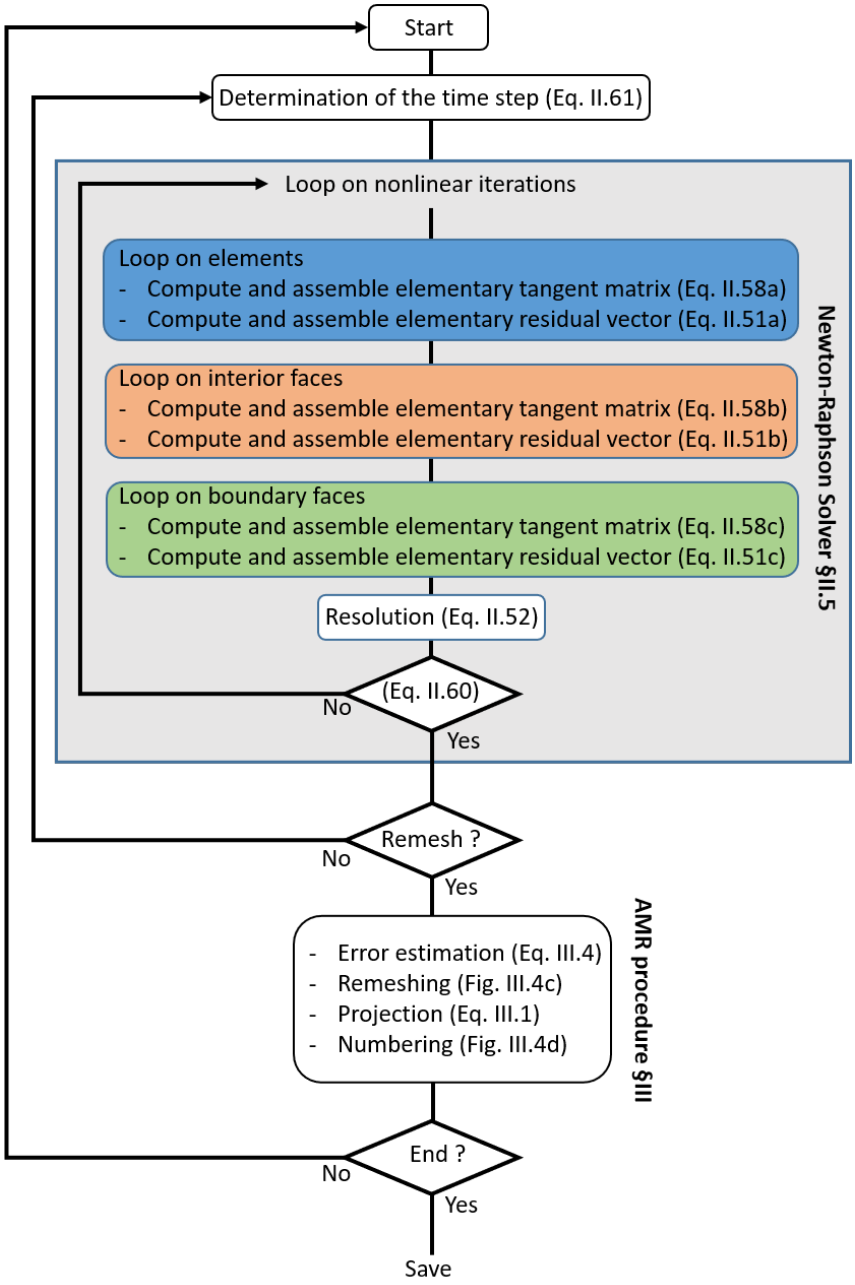


Figure II.5 – General architecture of the core algorithm for the numerical solving from *Rivage* code.

For each nonlinear iteration, the system has to be constructed and solved. The construction step is called the assembly algorithm and is detailed in the next paragraph. The solution step is carried by different solvers let to user choice. There are three possibilities:

- the in-house solver *SOLP* which is a direct solver using fully stored matrices. It is based on Gauss–Jordan elimination and can solve non-symmetric systems of linear equations. The algorithm can be found in [140];
- the in-house solver *SOLLC* which is a direct solver using skyline matrix storage (SKS) also known as band matrix storage. It is based on Gauss–Jordan elimination and can solve non-symmetric systems of linear equations. The algorithm can be found in [140];

- the free academic version 3.2.1 of *AGMG* which stands for *Iterative solution with AGgregation-based algebraic MultiGrid*. This software library solves large systems of linear equations with an aggregation-based algebraic multigrid method. It is known to perform well, in particular for scalar second order elliptic partial differential equations, but being purely algebraic makes it robust on a large range of problems like refined unstructured grids, convection-diffusion equations or problems with big jumps and large anisotropy. It uses matrices represented with the compressed sparse row (CSR) format. AGMG software and documentation were realised by Yvan Notay and can be found in http://agmg.eu. AGMG main algorithms are presented and justified in [141–143].

In this work, all numerical solutions were obtained using the direct methods because discrete systems remains usually small and we do not want to take care about convergence of iterative methods. Besides, we aim for the the parallelization of the solver in the future.

Assembly algorithm

From the notations and notions introduced in Section II.2.5, the system of Eq. (II.40) can be rewritten as a matrix system. It is assumed that local degrees of freedom are numbered contiguously within every element, so that basis functions and degrees of freedom can be reindexed:

$$\left\{\Phi_k^l : 1 \leq k \leq N_E, 1 \leq l \leq N_{\text{dof}}^E\right\} = \{\Phi_i : 1 \leq i \leq N_{\text{dof}}\}, \tag{II.62}$$

$$\left\{\chi_k^l : 1 \leq k \leq N_E, 1 \leq l \leq N_{\text{dof}}^E\right\} = \{\chi_i : 1 \leq i \leq N_{\text{dof}}\}. \tag{II.63}$$

Plugging the DG expansion Eq. (II.33) into the scheme Eq. (II.40), a nonlinear system of differential equations raises written with matrices which are defined such that, $\forall 1 \leq i, j \leq N_{\text{dof}}$:

$$M_{ij} = m(\Phi_j, \Phi_i), \tag{II.64}$$

$$A_{ij} = a(\Phi_j, \Phi_i), \tag{II.65}$$

$$L_i = l(\Phi_i, t). \tag{II.66}$$

where M, A and L are respectively called the mass matrix, the stiffness matrix and the load vector. These global matrices result from the assembly of local matrices corresponding to an element contribution. The stiffness matrix can be subdivided into three matrices with respect to volume, interior face and boundary face contributions highlighted by different colours in Eq. (II.28). The load vector might also be subdivided into two vectors: one for a boundary face contribution and one for a volume contribution if there is a source/sink term.

The assembly algorithm creates the matrix system to solve and consists in three loops: one loop for volume contributions, one for interior face contributions and one for boundary face contributions. For each loop, the local matrix is computed before being assembled into the corresponding global matrix. Then, the latter form the matrix system to solve. This process is illustrated in Fig. II.5. It is similar to assembly algorithm from FE methods except there are face contributions in addition. Details can be found in Rivière [26] or in Di Pietro and Ern [144].

Due to the local nature of element contributions, the global matrices are block-organized and mainly sparse. Diagonal blocks account for the element contribution while off-diagonal blocks account for interaction of elements through the face that connects them. Thus, numbering modifies directly the matrix structure since two neighbouring elements having distant numbering will produce far off-diagonal blocks. Reducing the bandwidth is important to decrease computational cost when specific storage are used for solvers. The standard numbering takes the element number progressively from

their construction. The Cuthill–McKee algorithm is particularly used to form a small bandwidth matrix [145]. However, *Rivage* employs Morton's code to renumber elements with Z-order space filling curves [146]. The method is very easy to implement, cost-effective and provide good numbering preserving locality even though the resulting connectivity matrix is not as optimal as for Cuthill–McKee algorithm. Effects of the reordering are observed in Fig. II.6 for mass matrix and tangent stiffness matrix taken from a simulation made with *Rivage*.

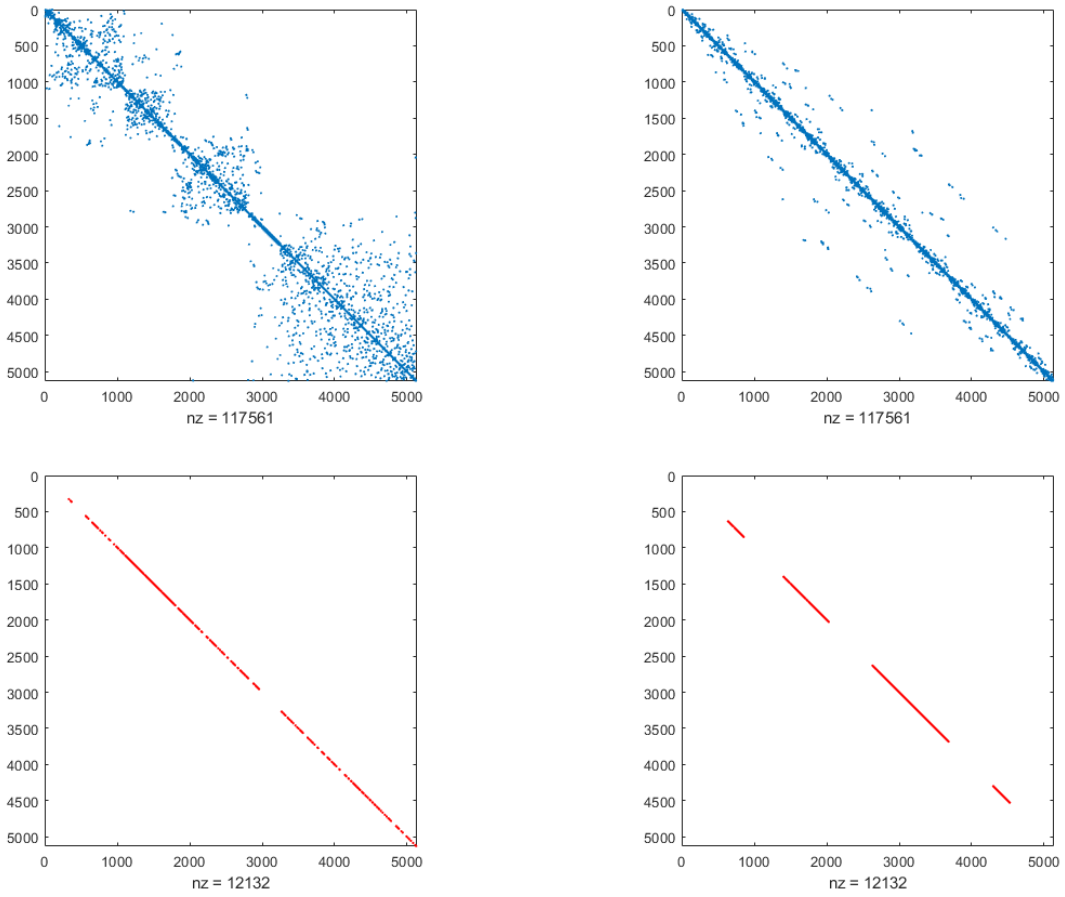


Figure II.6 – Sparsity patterns of the tangent stiffness matrix (top) and mass matrix (bottom) for standard numbering (left) and Morton-based numbering (right). *nz* stands for the number of non-zero values in the matrix.

II.7.2 Computing tools

Basis functions

Because there is no continuity constraint between elements for the test functions, their restriction to one element can be chosen independently from their restriction to other elements. Thereby, basis functions are specified to only one element which yields to a minimal stencil by reducing communications between elements. Besides, any polynomial functions basis is suitable which provides flexibility to DG methods. The choice of local basis is important because it affects the computational performance of DG methods, especially for high-order approximations. It exists two main families of basis functions: the nodal basis and the modal basis. The former give physical meaning to the degrees of freedom since they belong to nodes of elements such as for Fe methods. However, it is often diffi-

cult to have access to high-order because such basis require computations to find them. For instance, Lagrange or Hermite basis functions are nodal basis. One criterion to select them is the ease of integration. The modal basis are usually easy to implement for any polynomial degree. However, degrees of freedom have no meaning other than being coefficients in the DG expansion. Generally, the latter are preferred for DG methods because they can reach high-order accuracy easily. For instance, there are monomial, Taylor, Legendre or Dubiner basis. Two criteria are used to select them: orthogonality and hierarchism. If the basis is orthogonal or nearly orthogonal to respect of an inner product, this ensures that the condition number of the system does not grow too much with the polynomial degree and that some matrices are diagonal or nearly diagonal. A Gram–Schmidt procedure may be used for this purpose. If the basis is hierarchical, then, for a given polynomial degree p, the basis includes all the bases of polynomial degree $k < p$. This allows to effortlessly increase or decrease the polynomial degree locally.

In this work, a monomial basis is chosen because it is hierarchical and straightforward to implement to reach any order. The monomial basis principle is depicted in Fig. II.7. So, for this basis, it yields:

$$N_{\text{dof}}^{E} = \frac{1}{d!} \prod_{k=1}^{d} (p_E + k). \tag{II.67}$$

The Lagrange nodal basis – the classic FE basis – is also implemented in *Rivage* code, up to degree two for quadrangle and up to degree three for triangles. It can provide better approximations than monomial basis because it holds additional degrees of freedom for the same order, especially for quadrangles. Indeed, polynomial belongs to $\mathbb{Q}_p$ instead of $\mathbb{P}_p$. However, it is not hierarchical and reaching high-order is difficult. This is why the monomial basis is tensored to produce a new basis which is hierarchical and whose discrete space is enriched by tensor product, see Fig. II.7. The drawback is that computational cost increases due to higher dimension.

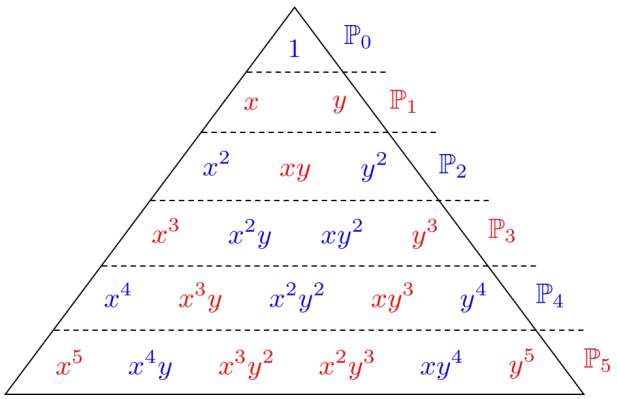


Figure II.7 – Principle of the construction of the basis functions based on the monomial basis. The triangle highlights the order hierarchy of the basis. Monomial basis corresponds to each stage while the tensor product of monomial basis corresponds to each coloured V-shaped mark.

Mapping and shape functions

Like for the FE methods, it is worth to use a mapping from a reference element to the – actual – physical element. Indeed, computing integrals over each element in the mesh would be costly because numerical quadratures should be adapted at each time. A preferred approach is to use change of variables to compute integrals over a fixed element. The procedure will be explained for 2D but can be extended to 3D. For this purpose, every element has vertices which are ordered in counter-clockwise.

Then, any element can be transformed from its corresponding reference element thanks to an affine transformation. In the case of *Rivage*, there are three reference elements: one for triangles, one for parallelograms and one for edges. Quadrangles are restricted to parallelograms because the mapping use an affine transformation which must preserve parallelism. For a given physical element E and its corresponding reference element $\hat{E}$, they have N_{nod} vertices (or nodes) whose 2D coordinates are respectively x_i, y_i and $\hat{x}_i, \hat{y}_i$, $i = 1, \ldots, N_{\text{nod}}$. There is a relation linking any point belonging to the physical element with the reference element:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 & \ldots & x_{N_{\text{nod}}} \\ y_1 & \ldots & y_{N_{\text{nod}}} \end{pmatrix} \begin{pmatrix} \hat{\varphi}_1(\hat{x}, \hat{y}) \\ \vdots \\ \hat{\varphi}_{N_{\text{nod}}}(\hat{x}, \hat{y}) \end{pmatrix}. \tag{II.68}$$

The functions $\hat{\varphi}_i, i = 1, \ldots, N_{\text{nod}}$ are called the shape functions. They are specific to reference element shape (triangle, parallelogram, *etc*) and defines the mapping. Generally, they are preferred to be polynomial and come from the FE methods. They match the Lagrangian basis functions, see [26, 140]. The reference element used in *Rivage* are depicted in Fig. II.8.

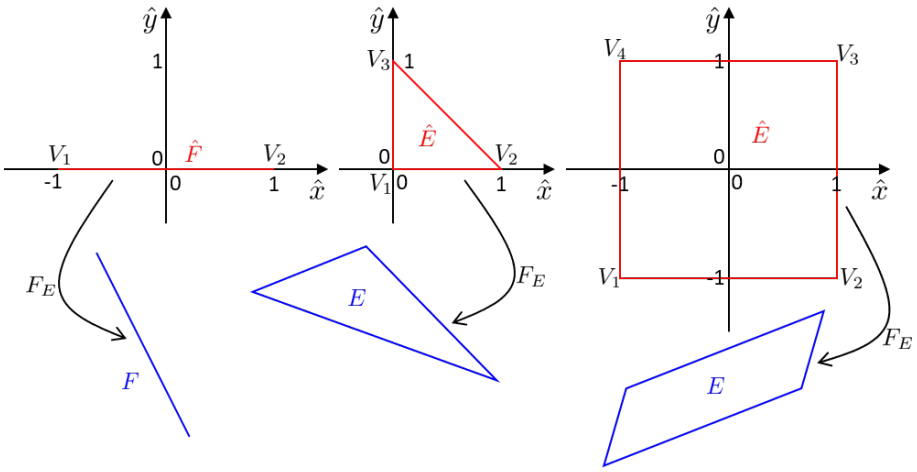


Figure II.8 – Three reference elements (red) with an example of related physical elements (blue): edge (left), triangle (middle) and parallelogram (right). V_i are the counter-clockwise vertices of the element and the origin O is at $(0, 0)$.

Formally, the mapping $F_E : \hat{E} \longrightarrow E$ can be rewritten as:

$$\begin{pmatrix} x \\ y \end{pmatrix} = F_E \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} \tag{II.69}$$

$$= J_E \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} + \begin{pmatrix} x_{\text{O}} \\ y_{\text{O}} \end{pmatrix}, \tag{II.70}$$

where $x_{\text{O}}, y_{\text{O}}$ are the coordinates of the physical element point related to the reference element origin point. The reference element origin point is set directly from the definition of the reference element. This is a predefined vertex for triangles and the barycentre for parallelograms. J_E is the Jacobian matrix

of the mapping F_E defined as:

$$J_E = \begin{pmatrix} x_1 & \dots & x_{N_{\text{nod}}} \\ y_1 & \dots & y_{N_{\text{nod}}} \end{pmatrix} \begin{pmatrix} \dfrac{\hat{\varphi}_1(\hat{x},\hat{y})}{\partial \hat{x}} & \dfrac{\hat{\varphi}_1(\hat{x},\hat{y})}{\partial \hat{y}} \\ \vdots & \vdots \\ \dfrac{\hat{\varphi}_{N_{\text{nod}}}(\hat{x},\hat{y})}{\partial \hat{x}} & \dfrac{\hat{\varphi}_{N_{\text{nod}}}(\hat{x},\hat{y})}{\partial \hat{y}} \end{pmatrix}. \tag{II.71}$$

If the determinant of J_E does not vanish, which is always satisfied when E is convex, there exists an inverse transformation ${F_E}^{-1} : E \longrightarrow \hat{E}$ such that:

$$\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = {F_E}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \tag{II.72}$$

$$= {J_E}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} - {J_E}^{-1} \begin{pmatrix} x_{\text{O}} \\ y_{\text{O}} \end{pmatrix}. \tag{II.73}$$

For a given function g, substitution theorem leads to a transformation of integral and the chain-rule gives transformation of derivatives:

$$\int_E g(x,y)\,\mathrm{d}E = \int_{\hat{E}} g \circ F_E(\hat{x},\hat{y}) |\det J_E|\,\mathrm{d}\hat{E}, \tag{II.74}$$

$$\int_E \nabla g(x,y)\,\mathrm{d}E = \int_{\hat{E}} {J_E}^{-\intercal} \hat{\nabla} g \circ F_E(\hat{x},\hat{y}) |\det J_E|\,\mathrm{d}\hat{E}. \tag{II.75}$$

When face integrals are considered, the reference face is taken to be the reference edge element (in 2D) as illustrated in Fig. II.8. For a given function g, the face integral is computed such that:

$$\int_F g\,\mathrm{d}F = \int_{\hat{F}} g \circ F_E(\hat{x},\hat{y}) \left| \det_F J_E \right| \mathrm{d}\hat{F}, \tag{II.76}$$

where F_E has to be understood as the mapping from the reference edge element (in 2D) to the physical face. The Jacobian matrix keeps the same definition. The last column of the Jacobian matrix will always hold zeros since face are $(d-1)$-objects. For this reason, the determinant is replaced with the following expressions from parametrization [140]:

$$\text{In 2D, } \det_F J_E \coloneqq \sqrt{{j_{11}}^2 + {j_{21}}^2}, \tag{II.77}$$

$$\text{In 3D, } \det_F J_E \coloneqq \sqrt{(j_{21}j_{32} - j_{31}j_{22})^2 + (j_{31}j_{12} - j_{11}j_{32})^2 + (j_{11}j_{22} - j_{21}j_{12})^2}, \tag{II.78}$$

where $j_{kl} = J_E(k,l)$.

Numerical integration

Quadrature rules will be used to calculate the numerical value of a definite integral. One particularity of DG methods is that integrals need to be evaluated exactly both over edges and elements, and for polynomial functions of various degrees, possibly high. There are many possible quadrature rules with their own advantages and drawbacks. The Gaussian quadrature rules is considered in this work because it is widespread and presents a good compromise between rule difficulty and accuracy. For edges in 1D, the rule is given by:

$$\int_{-1}^{1} f(x)\,\mathrm{d}x = \sum_{i=1}^{N_G} w_i f(x_i), \tag{II.79}$$

where w_i and x_i for $i = 1, \ldots, N_G$ are respectively the Gauss quadrature weights and nodes. Their values can be found in many textbooks, for example in [140]. General formulas are available but it is preferred to store values directly in the code to facilitate access during computation since this type of calculation is performed many times during simulation. Equation (II.79) is exact for polynomials of degree $2N_G - 1$ or less and is sometimes referred as the Gauss-Legendre quadrature. Since DG methods reach easily high-order approximation, it is important to get high-order quadrature rules.

For parallelograms in 2D, the one dimensional Gauss quadrature rule is applied in each direction:

$$\int_{-1}^{1}\int_{-1}^{1} f(x)\,\mathrm{d}x = \sum_{i=1}^{N_G}\sum_{j=1}^{N_G} w_i w_j f(x_i, x_j). \tag{II.80}$$

For triangles in 2D, the Gauss quadrature rule takes specific set of values.

Summary

Since the solution of Richards' equation is not available analytically for most cases, numerical methods are necessary. DG methods based on penalty are proposed in this study because they show useful mathematical properties. Indeed, they rely on a element-wise weak formulation which turns to be flexible, both mathematically and numerically. This allows to handle complexed modelling features and advanced computational methods. In particular, the treatment of seepage boundary condition or dynamic forcing condition is natural since they are directly incorporated into the weak formulation thanks to penalization. Moreover, penalty and symmetry terms provide stability and continuity at different level of control following the variants of the formulation. Finally, the lack of continuity constraint between elements gives the possibility to use different polynomial approximations and non-conforming mesh which are the basis properties to reach high-order approximation and use adaptive mesh refinement. The scheme is mass conservative at machine precision thanks to natural local mass balance.
Implicit schemes are chosen for time discretization to guarantee stability. Backward difference formula methods are chosen because they are known to be good to deal with stiff equations and they provide high-order time scheme which is convenient if spatial errors are greatly reduced.
Due to nonlinearities present in water content and hydraulic conductivity, solving Richards' equation requires a robust iterative process. Even though the linearization by the classic Newton-Raphson method is quadratic convergent, it suffers from reliability because of poor initial guess or singular Jacobian matrix. The fixed-point method is preferred because it is more robust despite being linearly convergent. Robustness is strengthened thanks to an heuristic adaptive time stepping based on nonlinear iterations.
The numerical methods were implemented in Fortran from scratch to build a computational code named *Rivage*. This program was thought to be as general as possible to handle a large range of user choices: non-conforming/unstructured/hybrid mesh of triangles/quadrangles, solvers based on different matrix storage, high-order approximation, several basis functions, *etc*. Future extensions should be done naturally like parallelizing and 3D.

Chapter III

Adaptive strategy based on discontinuous Galerkin method

"There is nothing like looking, if you want to find something. You certainly usually find something, if you look, but it is not always quite the something you were after."
— J. R. R. Tolkien, *The Hobbit, or There and Back Again*, 1937

One of the main characteristics and quality of the DG method is to facilitate mesh adaptation by element refinement and coarsening (h-adaptation), as well as by enrichment of approximation order (p-adaptation). Combination of this two processes is called hp-adaptation.

In numerical simulation, mesh adaptation is an essential tool to get accuracy and cost-effectiveness. Indeed, local mesh modification can serve as a capturing technique to resolve physical phenomena or as a control technique to monitor numerical errors. Adaptation strategy is based on two key ingredients defining main goals of the mesh adaptation. Firstly, the overall process has to be chosen to target goals of adaptation. It is sometimes referred as adaptivity. Many different techniques exist to optimize

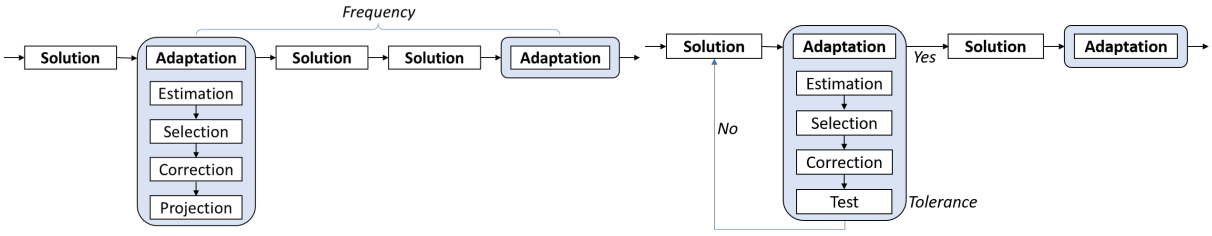
the mesh, select elements to adapt and refine/coarsen these elements [113]. Secondly, an error criterion has to be decided in order to localize zones to adapt. One possibility is to use an error indicator which defines a quantity of interest taken to be typical of the physics, very other related to solution gradient [147]. Another possibility is to use error estimator which comes from theoretical derivations of the numerical scheme [148, 149]. Error estimation is said to be *a priori*, if estimate is established before the solution is known to exist, or *a posteriori*, if estimate uses the already computed solution. Besides, *a posteriori* estimation can be either explicit or implicit [150]. Estimator is explicit when it can be computed directly from the numerical solution and the problem parameters, or implicit when its computation needs an additional problem to solve. Lots of *a posteriori* error estimations are found in literature [147, 149, 151] and they are sometimes classified according to their underlying principle. For instance, they are said to be hierarchical [152–154] when it measures the difference between the numerical solution and a reference numerical solution obtained with an enhanced scheme like a more refined region [155, 156]. Others are based on residuals [157], flux reconstruction [158], dual spaces [159], *etc*.

The strategy presented in this chapter aims to optimize a compromise between feasibility, relevance, accuracy and computation time. This is a block-based mesh refinement which is not applied systematically to save computational time. Then, two criteria are introduced: one is an error indicator and the other is an estimation-based error indicator. The latter is called in this way because no theoretical derivation supports it but rather an heuristic comparison from other error estimators. Finally, some variations on the numerical scheme are devised. They are taking the most of the DG flexibility to propose improvements. One is the use of the weighted DG method which relaxes average and penalty definitions, and the other is an extrapenalty term for the jumps of total flux (derivatives).

III.1 Adaptive mesh refinement

III.1.1 Adaptivity

In this part, the *h*-adaptation tool that has been developed is presented. Mesh refinement consists in concentrating the discretization points (element, volume, ...) at the places where it is necessary, in a sense that must be defined. Adaptive Mesh Refinement techniques (AMR) are now widely used and have since proven their efficiency. As said earlier, mesh refinement may target different goals and so use different strategies defining adaptivity. To get an idea of the process for *a posteriori* estimation, several steps can be sketched, see Fig. III.1.



(a) Sequential adaptivity (b) Iterative adaptivity

Figure III.1 – Different types of adaptation strategy

First of all, the problem is solved to get the numerical solution (**solution**). Then, error estimates are computed for each element (**estimation**). Next, elements are eventually marked for refinement/coarsening (**selection**). Last, the mesh is adapted accordingly (**correction**). Each of this step can

be performed by many different techniques: evaluation is proper to the numerical method, estimation is based on the chosen error indicator/estimator, selection may be proportional or absolute, user-defined or automatic and correction depends on the element geometry and mesh structure. In addition, Fig. III.1 illustrates the two main adaptive approaches. Sequential adaptivity is done throughout the simulation when it is considered needed whereas iterative adaptivity is done systematically for every computed solution. Usually, sequential adaptation wants to stay efficient in term of computational time and iterative adaptation wants to reach optimality by finding the mesh minimizing numerical errors.

AMR techniques are able to work with many different meshes: on 2D Cartesian grids for quadtree [160–163], on 3D Cartesian grids for octree [164, 165], and on anisotropic mesh [166–168]. As shown in Fig. III.2 (right), anisotropic mesh refinement distorts the mesh to concentrate the elements where they are needed. They are sometimes called *r*-adaptation or relocation refinement and belong to moving mesh methods. This method is used for stabilized finite elements on any conforming mesh and it allows very local refinement, but at the price of frequent remeshing in order to follow the physics of the phenomenon. Examples for porous media flow include [108, 169, 170]. The refinement by quadtree or octree is mainly used for finite volumes on structured but non-conforming Cartesian mesh. Very local refinement is also possible. The tree-like structure of the mesh data allows, through appropriate algorithms, a quick access to a cell or its neighbours.

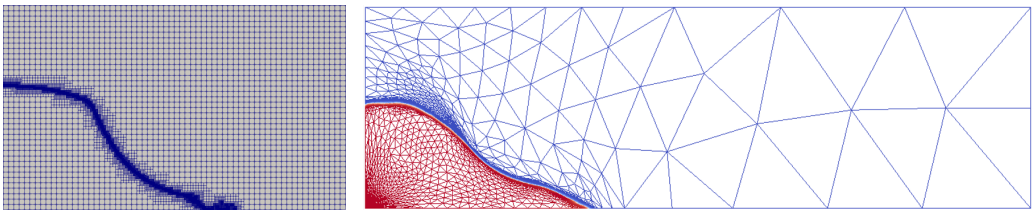


Figure III.2 – Example of mesh refinement applied dam break benchmark. Using Basilisk code with quadtree and finite volume (left) and using finite element and anisotropic mesh adaptation [168] (right).

III.1.2 Block-based structure

There is a number of AMR types that have been developed [171–173], see Fig. III.3. In cell-based AMR, refinement occurs element-wise and is generally built upon a parent-child data structure. Cell-based AMR would be interesting since DG methods provide local features suitable for such AMR. Patch-based AMR groups a certain number of elements into a patch where the requested refinement is done, providing regular grids. Finally, block-based AMR predefines coarse elements in which a refinement is performed. Block-based is suitable for structured Cartesian mesh.

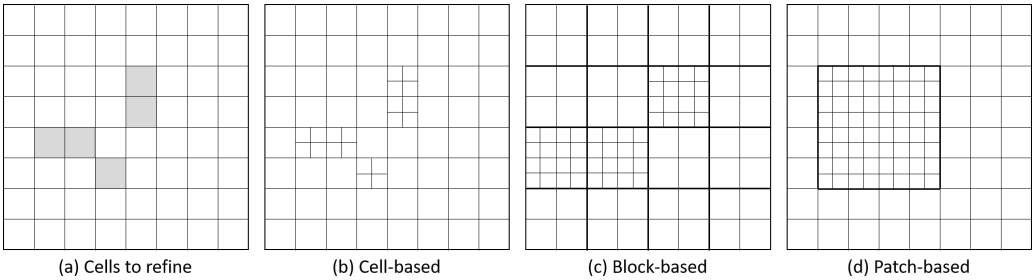


Figure III.3 – Cell-, block- and patch-based AMR techniques for Cartesian grids

In this work, applications with complex geometries are aimed, or later, couplings of models on hybrid mesh. This is why a non-conforming unstructured mesh is used like in [174]. Although computing time may be lost compared to fully structured Cartesian codes, block-based AMR is adopted for several reasons. First, block-based mesh can be viewed as a macroscale mesh surrounding large regions of the domain. Provided that physical phenomena like moving fronts stay in the block, the block refinement is kept avoiding frequent remeshing. This idea follows some remarks found in [20]. Secondly, block-based structure provides a basis to parallelize the code since only inter-block connectivity is required. Finally, block-based structure leads to solving techniques like local time stepping or domain decomposition which can be useful for Richards' equation [175].

III.1.3 Selection procedure and corrective mechanism

In accordance with the simulations presented in this work, only the bidimensional case is presented. The strategy adopted is then as follows and illustrated in the Fig. III.4.

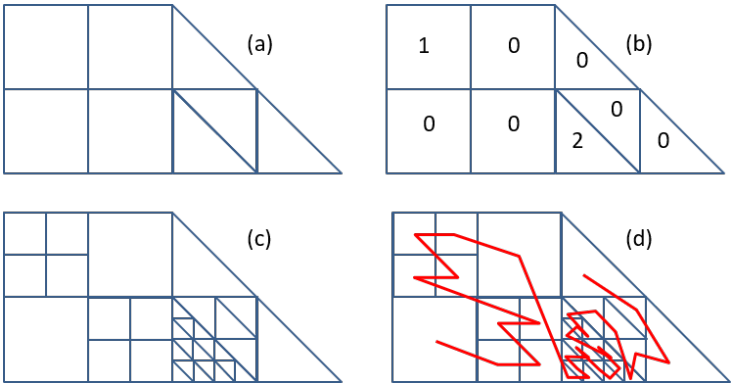


Figure III.4 – AMR Strategy: (a) Block mesh; (b) Level of mesh refinement; (c) mesh generation; (d) Morton numbering.

The unstructured mesh is composed of quadrilateral and/or triangular elements, where each element, as in [174], defines a *root element* or *block*, see Fig. III.4(a). Each of this block can be refined appropriately into an isotropic grid of elements. All elements generated by the blocks are gathered into a list. It is not a parent-child data structure element-wise but blocks know their elements likewise.

Mesh refinement level is defined for each block, initially by the user and thereafter in accordance with the chosen mesh refinement criterion, see Fig. III.4(b). Coarsening and refinement are decided according to threshold values $0 < \beta_c < \beta_r$. If η denotes the error estimate, three situations are encountered:

- For $\beta_c < \eta < \beta_r$, the block remains unchanged and so for the related elements;
- For $\beta_r < \eta$, the block is refined so its elements are split into four isotropic subelements;
- For $\eta < \beta_c$, the block is coarsened so four neighbouring subelements are merged into one element.

Limitations of such selection are described in [176] as well as a procedure to overcome them. The mesh refinement level is adapted according to the rule that the ratio of mesh refinement should not exceed 2 between two neighbouring blocks producing a graded mesh. As illustrated in [177], this constraint allows smooth transitions between refined and unrefined regions.

The mesh is then built using a quadtree graph (available for quadrangle and triangle) in order to define easily and precisely the neighbouring elements of each face. This procedure is pursued until the desired level of mesh refinement is reached as shown in Fig. III.4(c). In the case of evolutionary problems, this implies frequent remeshing to follow the studied phenomenon. This is why, we prefer to widen the stencil with a coarse root mesh, namely blocks, and therefore to use the whole quadtree graph, in order to remesh less often. It is illustrated for quadrangles and triangles in Fig. III.5. Finally, a space filing curve using Morton numbering (Z-order) is built in order to number easily the degrees of freedom line in Fig. III.4(d).

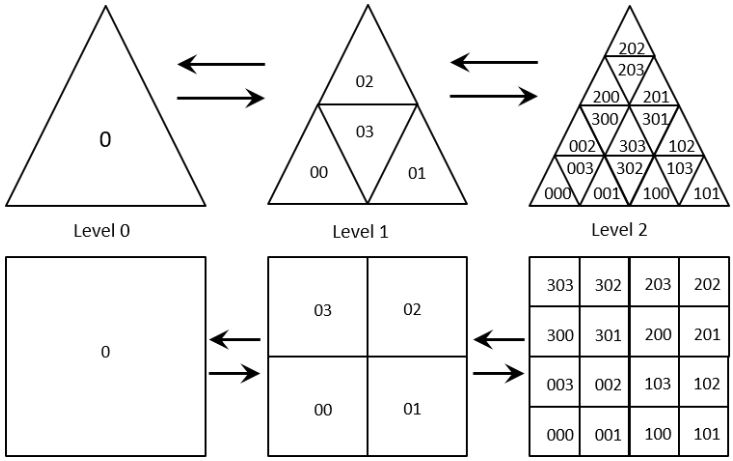


Figure III.5 – Mesh refinement for quadrangle and triangle using quadtree.

III.1.4 Projection by local problem

During a simulation, if a mesh refinement or coarsening process takes place, a new evaluation of the unknown field must be carried out to keep solution information. The prolongation (refinement) and restriction (coarsening) process are facilitated by the fact that the quadtree graph connect explicitly the "mother" element to its four "daughter" elements. To do that, one may use local interpolation/extrapolation. This is not straightforward in practice because it requires to identify the number of points and their position on the element to construct a right-determined system giving a fairly close solution. In the present study, the projection is made by solving a local DG problem. This technique gives good degrees of freedom by performing calculations at quadrature points. The element-wise weak formulation corresponding to $h_{E_{new}} = h_{E_{\text{old}}}$ reads:

$$\int_{E_{\text{new}}} h_{E_{\text{new}}} v \, \mathrm{d}E_{\text{new}} = \int_{E_{\text{new}}} h_{E_{\text{old}}} v \, \mathrm{d}E_{\text{old}}. \tag{III.1}$$

Then, mass matrix can be used to project the solution on the new mesh. This method is also used to enforce the initial condition $h = h_0$, see Section II.3.3.

III.2 *A posteriori* error estimation

III.2.1 Gradient-based error indicator

An *a posteriori* error indicator is employed to show how a simple physical quantity is able to provide information about refinement. This error indicator is inspired from Miller *et al.* [20] where they

use changes in effective saturation. Here, the error indicator is based directly on changes in the solution that is to say the gradient of hydraulic head h. This allows to work both for the saturated and unsaturated regions. So, it writes:

$$C_E = \frac{1}{\max\limits_{E\in\mathcal{E}_h}\left(\frac{1}{|E|}\|\nabla h\|_{L^2(E)}\right)}\frac{1}{|E|}\|\nabla h\|_{L^2(E)}. \tag{III.2}$$

The error indicator is normalized so that $0 \leq C_E \leq 1$.

III.2.2 Estimation-based error indicator

Adaptive refinement aims to control numerical approximation within a prescribed tolerance employing the least computational resources. That is why it is essential to have an *a posteriori* error estimation which gives information about the true error only from the computed numerical simulation and the given data of the problem. The true error for a function u is defined as:

$$e_h^n \coloneqq u - u_h^n \tag{III.3}$$

The computation of the *a posteriori* error estimation should be as economical as possible compared to the numerical solution. Besides, other features are desired for the error estimator:

- locality, namely, the error estimator gives information element-wise about the error distribution which is relevant for local adaptation;
- reliability, namely, the error estimator gives an upper bound on the true error up to a multiplicative constant. This ensures that the true error is below the tolerance prescribed for the error estimator;
- efficiency, namely, the error estimator gives a lower bound on the the error up to another multiplicative constant. This ensures that the true error is not too far from the error estimator to achieve the prescribed tolerance cost-effectively.

The product of the two aforementioned constants measures the error estimator quality analogously to a condition number. This quantity gives details about error estimation calibration in the sense that it should stay bounded regardless of the parameters involved in the differential equation [151]. Another powerful property is the derivation of a guaranteed *a posteriori* error estimate where guaranteed means the upper bound of the error is explicitly computable that is to say the reliability constant is guaranteed to be one. Ern and Vohralík achieved this goal using divergence-conforming flux reconstruction for mixed *h*-adaptive DG methods, see for example [178] for convection-diffusion-reaction problems. Even though residual-based error estimation overestimates the true error by a significant factor, this approach is interesting because it is simple and well-developed for lots of problems, including time or *hp*-adaption.

Another *a posteriori* error indicator is introduced but it is based on error estimation. Indeed, Verfürth [151] developed *a posteriori* error estimators based on residual and derived thanks to suitable norms in the context of finite elements methods for many equations. These estimates are valid for many equations and he worked in particular on nonlinear parabolic problems such as Richards' equation. For this kind of problem, Verfürth [179] used general variational approach based on the concept of very weak solutions to prove reliability as well as global and local efficiency for non-standard time discretization such as implicit Runge-Kutta schemes. However, this analysis is sophisticated and subject to step

size restriction in space and time. Then, he removed this constraint in [180, 181] for standard time discretization, ranging from implicit Euler scheme to Crank-Nicolson scheme and midpoint rule, thanks to energy norm augmented by dual norm of time derivative. It results in truly reliability and efficiency, provided that solving an additional discrete problem is agreed. Later on, this type of residual-based derivation was adapted by Melenk and Wohlmuth [182] to *hp*-adaptive finite elements methods for elliptic problems generalizing the Clément/Scott-Zhang interpolation operator. Afterwards, Houston *et al*. [183] developed residual-based *a posteriori* error estimation for *hp*-adaptive interior penalty DG methods applied to elliptic problems exploiting the DG energy norm. Finally, Schötzau and Zhu [122, 184] extended this approach to convection-diffusion equations both for non- and *hp*-adaptive DG methods.

Greatly inspired by these works and supported by an heuristic analysis, the following *a posteriori* estimation-based error indicator is used:

$$\eta_E^{n\,2} = \eta_{\mathrm{R}E}^{n\,2} + \eta_{\mathrm{F}E}^{n\,2} + \eta_{\mathrm{J}E}^{n\,2}, \tag{III.4}$$

where the element residual, the face residual and the solution jump sub-estimates are respectively:

$$\eta_{\mathrm{R}E}^{n\,2} = \frac{h_E^2}{p_E^2 \lambda_{\mathrm{m}}(\mathbb{K})} \left\| \frac{\theta\left(u_h^{n+1}\right) - \theta(u_h^n)}{\tau^n} - \nabla \cdot \left(\mathbb{K}\left(u_h^{n+1}\right) \nabla u_h^{n+1} \right) \right\|_{L^2(E)}^2, \tag{III.5}$$

$$\begin{aligned} \eta_{\mathrm{F}E}^{n\,2} = & \sum_{F \in \partial E \cap \mathcal{F}_h^{\mathrm{I}}} \frac{h_F}{2 p_F \kappa_{\mathrm{m}}} \left\| \left[\!\left[\mathbb{K}\left(u_h^{n+1}\right) \nabla u_h^{n+1} \cdot \boldsymbol{n} \right]\!\right] \right\|_{L^2(F)}^2 \\ & + \sum_{F \in \partial E \cap \mathcal{F}_h^{\mathrm{N}}} \frac{h_F}{p_F \kappa_{\mathrm{l}}} \left\| q_{\mathrm{N}} - \mathbb{K}\left(u_h^{n+1}\right) \nabla u_h^{n+1} \cdot \boldsymbol{n} \right\|_{L^2(F)}^2, \end{aligned} \tag{III.6}$$

$$\begin{aligned} \eta_{\mathrm{J}E}^{n\,2} = & \sum_{F \in \partial E \cap \mathcal{F}_h^{\mathrm{I}}} \frac{1}{2} \left(\varrho_F^{\mathrm{I}} + \frac{h_F}{p_F \kappa_{\mathrm{m}}} \right) \left\| \left[\!\left[u_h^{n+1} \right]\!\right] \right\|_{L^2(F)}^2 \\ & + \sum_{F \in \partial E \cap \mathcal{F}_h^{\mathrm{D}}} \left(\varrho_F^{\mathrm{D}} + \frac{h_F}{p_F \kappa_{\mathrm{l}}} \right) \left\| u_{\mathrm{D}} - u_h^{n+1} \right\|_{L^2(F)}^2. \end{aligned} \tag{III.7}$$

The quantities $\lambda_{\mathrm{m}}(\mathbb{K})$ and $\lambda_{\mathrm{M}}(\mathbb{K})$ stand respectively for the lowest and largest eigenvalue of $\mathbb{K}$ on E. For a face F of normal $\boldsymbol{n}$, the minimum of normal component from neighbouring $\mathbb{K}$ is chosen: $\kappa_{\mathrm{m}} \coloneqq \min(\kappa_{\mathrm{l}}, \kappa_{\mathrm{r}})$ where $\kappa = \boldsymbol{n}^{\intercal} \cdot \mathbb{K}\left(u_h^{n+1}\right) \cdot \boldsymbol{n}$. Here, the time discretization is based on the implicit Euler scheme but the *a posteriori* estimation-based error indicator can be determined for other time schemes such as in [179]. Global estimates are obtained with:

$$\eta^{n2} = \sum_{E \in \mathcal{E}_h^n} \eta_E^{n\,2}. \tag{III.8}$$

Unlike the aforementioned papers, there is no rigorous mathematical proof for this residual-based energy norm *a posteriori* error estimation. Extension to the case of nonlinear and time-dependant parabolic equation solved by *hp*-adaptive DG methods remains an open problem. However, some remarks can be drawn. Firstly, the numerical treatment of the time dependence and of the nonlinearity resorts to consider a sequence of linear and steady problems. Secondly, as noted by Verfürth [151], the element residual term is related to the residual of the numerical solution with respect to the strong form of the equation which may be viewed as the error from the nonlinear process. The face residual term is related to the boundary operator which is associated with the strong and weak forms

of the differential equation. It reflects, on one hand, that the numerical solution gradient – the flux – is discontinuous and on the other hand, that Neumann boundary conditions may not be satisfied. The solution jump term is related to the penalization which are associated with the weak form of the equation. It reflects that the numerical solution is discontinuous at interior faces in the DG framework and Dirichlet boundary conditions may be slightly violated according to the penalty coefficients. Moreover, the second term and third parts of the estimation-based error indicators measure how valid the hypotheses on the seek solution are to derive the DG formulation.

III.3 Variation on the numerical scheme

III.3.1 Weighted discontinuous Galerkin framework

In the context of convection-diffusion equations or coupling of parabolic/hyperbolic domains, the solution may hold spurious oscillations at internal layers for vanishing or varying diffusion [185]. This situation is typical of porous media problems where there are material heterogeneity and degenerate diffusion. Even in the case where diffusion is nonzero but small, the numerical solution imitates this limiting case. One interpretation is the flux stays continuous but the solution mimics a discontinuity. Then, even though the mathematical description has some regularity, the numerical scheme is unable to recognize the sharp internal layer leading to numerical instabilities. This can trouble mass conservation.

Even though this phenomenon may be resolved by mesh refinement, the corresponding computational cost is substantial to maintain a physically acceptable solution. Slope limiters are another possibility to control these undershoots/overshoots but further computational developments are needed to cope with the geometry of elements in high-dimension [26, 84]. A satisfactory approach would be to design a DG method that can handle internal layers in an automated fashion. That is the reason why the weighted DG framework was introduced. Following earlier works [186–188], Proft and Ern proposed one possible remedy in [185]. This work was continued by Proft and Rivière on one hand [26, 34, 35], and by Ern, Di Pietro, Zunino, Stephansen, Guermond on the other hand [32, 33, 144, 178]. Proft and Rivière called their two proposed methods *improved* and *adapted* DG methods where they focus on exploring different average flux definitions. Nevertheless, their methods can be viewed as particular cases from Ern-Di Pietro work referred as the *weighted* discontinuous Galerkin methods (WDG).

WDG methods work thanks to two key ingredients which can be seen as a way to incorporate into the definition of the scheme some partial knowledge of the solution. The first one is the use of weighted averages instead of the standard arithmetic average inside the discontinuous formulation. The amount of diffusive flux for each side of a face is controlled and is used to formulate the consistency terms in the method analysis. The second ingredient is to relax the face penalization used for continuity constraint by a coefficient depending on the harmonic mean of the diffusivity of the neighbouring elements. Such a penalty strategy turns out to tune automatically the amount of local penalty to regulate the degree of smoothness of the approximate solution. Analysis and details can be found in [26, 32] and related. One should note that weighted framework was developed for convection-diffusion equation and not parabolic equation like Richards' equation. However, Proft and Ern demonstrate how it is useful even for purely parabolic problems when the diffusion coefficient varies spatially [185].

Internal layers depend on the spectral structure of the diffusivity, especially the degree of anisotropy is related to the condition number. The hydraulic conductivity is a matrix-valued function which

corresponds to an anisotropic diffusion. Its degree of anisotropy is $\Delta_{\mathbb{K}} \coloneqq \dfrac{\lambda_{\mathrm{M}}(\mathbb{K})}{\lambda_{\mathrm{m}}(\mathbb{K})}$. Weights across a face are positive numbers such that:

$$\omega_{\mathrm{l}} + \omega_{\mathrm{r}} \coloneqq 1. \tag{III.9}$$

The weighted average and the conjugate weighted average are respectively defined as:

$$\{\!\{u\}\!\}_{\omega} \coloneqq \omega_{\mathrm{l}} u_{\mathrm{l}} + \omega_{\mathrm{r}} u_{\mathrm{r}}, \tag{III.10}$$

$$\{\!\{u\}\!\}_{\overline{\omega}} \coloneqq \omega_{\mathrm{r}} u_{\mathrm{l}} + \omega_{\mathrm{l}} u_{\mathrm{r}}. \tag{III.11}$$

Then, the relation $\forall u, v \in S_p(\mathcal{E}_h^n), [\![uv]\!] = \{\!\{u\}\!\}_{\omega} [\![v]\!] + [\![u]\!] \{\!\{v\}\!\}_{\overline{\omega}}$ is replacing Eq. (II.19) in the weak formulation derivation. One interpretation of this weighted average is that the mortar between elements is proportional to the stiffness of two springs in series. This choice of average is employed to find an equivalent hydraulic conductivity for Richards' equation, see [89]. Moreover, the weighted average defines a numerical flux which is related to the ones from the Local Discontinuous Galerkin method which is known to better behave with discontinuous solutions.

The weighted diffusion penalty coefficient γ_F is taken to be the weighted average of diffusivities which results to be equivalent to the harmonic mean of the neighbouring normal hydraulic conductivities across the face [32]:

$$\gamma_F = \{\!\{\kappa\}\!\}_{\omega} = \frac{2\kappa_{\mathrm{l}}\kappa_{\mathrm{r}}}{\kappa_{\mathrm{l}} + \kappa_{\mathrm{r}}}. \tag{III.12}$$

Standard DG methods are designed to weakly enforce the continuity of the numerical solution between mesh elements. However, excessive penalty results in spurious oscillations for very small and vanishing diffusivity. Allowing stronger discontinuities leads to smaller overshoots/undershoots and reduces the spurious oscillations in the vicinity of the internal layer by preventing their propagation. The stability of the scheme is improved. It is interesting to note that several authors consider a specific definition for the penalty parameter to adjust the size of the jumps according to problem variables, space dimension, approximation order and element shape, see [29, 109, 189, 190] where penalty treats problems with high heterogeneity or discontinuity.

Standard DG methods assume $\omega_{\mathrm{l}} = \omega_{\mathrm{r}} = \dfrac{1}{2}$, which reduces the weighted average to the arithmetic average, and $\gamma_F = 1$. The WDG method chooses other definitions. Following definitions in [32], let the weights be:

$$\begin{cases} \omega_{\mathrm{l}} = \dfrac{\kappa_{\mathrm{r}}}{\kappa_{\mathrm{l}} + \kappa_{\mathrm{r}}}, \quad \omega_{\mathrm{r}} = \dfrac{\kappa_{\mathrm{l}}}{\kappa_{\mathrm{l}} + \kappa_{\mathrm{r}}} & \text{if } \kappa_{\mathrm{l}} + \kappa_{\mathrm{r}} \neq 0, \\ \omega_{\mathrm{l}} = \omega_{\mathrm{r}} = \dfrac{1}{2} & \text{otherwise.} \end{cases} \tag{III.13}$$

WDG methods assume that diffusivity discontinuities fit the mesh [32]. This is generally the case for heterogeneous media but it is not true for nonlinear diffusivity such as in Richards' equation. Then, sharp internal layers move dynamically without *a priori* knowledge of the location and may occur inside an element. So the weighted framework is not expected to work in this situation. Nevertheless, a suitable AMR as described previously can be used as a capturing technique. This idea is one of the main novelties proposed in this thesis. The weighted framework and mesh adaptation are working dynamically in synergy: the former changes the smoothness nature of the numerical solution while the latter tracks the internal layer linked to nonlinearity thanks to refinement driven by the *a posteriori* estimation-based error indicator of Eq. (III.4).

III.3.2 Extrapenalization

Rivière introduced an additional penalty term in the DG formulation [26]:

$$J_{\mathrm{X}}(h,v) = \sum_{F\in\mathcal{F}_h^{\mathrm{I}}} \int_F \varrho_F^{\mathrm{X}} [\![\mathbb{K}(h-z)\nabla h\cdot\boldsymbol{n}]\!][\![\mathbb{K}(v-z)\nabla v\cdot\boldsymbol{n}]\!]\,\mathrm{d}F \tag{III.14}$$

It constrains the jump of derivatives of the function to be zero. Therefore, it weakly imposes the continuity of the numerical total flux which is wanted from a physical point of view but also mathematically to be able to derive the weak formulation by subtracting it Eq. (II.18). That is why it is referred as an extra stabilization term. Nevertheless, the numerical analysis of the method is independent of this term and so Rivière discarded it without further use or explanation.

The total flux regularity comes into play to explain the presence of wetting front/internal layer due to nonlinearity and heterogeneity. The total flux becomes discontinuous if sharp wetting fronts produce spurious oscillations which is not desirable. As a result, an extrapenalty term is interesting for Richards' equation to preserve continuity of the Darcian flux. Moreover, this extrapenalty term controls the jump of the flux which is measured by one part of the *a posteriori* error estimator, see Eq. (III.6) and so have influence on mesh adaptation.

Summary

DG methods offer great abilities to make mesh adaptation by working with local approximation. This enables non-conforming mesh and element-wise different basis functions, resulting in the so-called *h*- and *p*-adaptation. This thesis focuses on *h*-adaptation by proposing an adaptive strategy whose goals are to minimize computational cost and reaching accuracy by capturing physical phenomena of Richards' equation.
A block-based AMR is chosen because it allows to save computational time. First, blocks partition the domain into large regions where the solution is expected to behave similarly. Internal layers due to heterogeneous diffusion are attached to blocks interface while moving fronts propagate through the domain but should stay within a block for a while. Then, if the front is well resolved by the block discretization, less frequent remeshing is needed so computational time is saved. Second, the block-based structure leads to other advanced solving making domain decomposition and parallelization easier.
Two estimates are proposed to drive mesh adaptation which is used as a capturing technique for sharp fronts rather than a control technique to reduce numerical errors. With this in mind, an error indicator based on hydraulic head gradients is proposed as well as an error indicator based on *a posteriori* error estimation. Their performance will be evaluated in the next chapter for Haverkamp's test-case and Tracy's benchmark.
Taking the most of DG flexibility, two modifications are also proposed for the numerical scheme. One is to use the weighted DG method which has demonstrated efficiency in the context of convection-diffusion equation to resolve sharp internal layers due to heterogeneous and vanishing diffusivity. A diffusivity-based weighted framework is applied to flux average and penalty definition. This allows for jumps in the solution which are expected to help numerical approximation to better resolve sharp fronts. Besides, thanks to weak penalization, an extrapenalty term is introduced to control jump of derivatives and guarantee total flux continuity which is important both for the numerical scheme and Richards' equation behaviour.

Chapter IV

Validation and numerical results

"We are reaching the stage where the problems that we must solve are going to become insoluble without computers. I do not fear computers. I fear the lack of them."

— Isaac Asimov, *The Computer Society: The Age of Miracle Chips (Time)*, 1978

Following the numerical methods presented in the previous chapters, the *Rivage* code and the adaptive strategy are validated against test-cases and experimental comparisons.

First of all, few numerical experiments are carried out with artificial cases built on the analytical solution of Poisson's equation. They allow to get some basic behaviours of stability and convergence for the DG methods.

Then, a series of test-cases are used to validate the implementation of the numerical methods. All of them are 1D infiltrations into a soil column but they hold different hydraulic properties and various boundary conditions. The solution of Richards' equation is more or less difficult to get and these

cases are found largely in the literature, providing reference numerical solutions. That is why these infiltration test-cases are good points of comparison to validate the solution of Richards' equation by the newly developed code. In particular, Polmann's test-case illustrates the importance of mesh resolution to get oscillation-free solutions of Richards' equation. In this case, the WDG method can be useful to improve the approximation quality.

Finally, performance of the adaptive strategy proposed in this thesis is demonstrated. A 2D benchmark with a known exact solution for Richards' equation is employed to show convergence and the error estimation efficiency to drive mesh adaptation. Laboratory experiments involving 2D water table control, as well as their related simulations available in the literature, serve as a keystone to know if the method is able to simulate complex porous media flow dynamics with recharge, drainage and seepage. Finally, experimental data from the reservoir impoundment of La Verne dam are used to assess the accuracy and the robustness of the method. This is a demanding case because there are several materials, stiff hydraulic properties, dynamic boundary condition and seepage. The results of the simulation are discussed to highlight the potential of the solving strategy.

IV.1 Numerical experiments

These test-cases aim to highlight some properties of DG methods for simple problems whose analytical solution is known. They are built upon the Poisson's equation. In particular, convergence and stability tests are targeted.

IV.1.1 One-dimensional analytical tests

For this first test-case, theoretical convergence of the DG methods are checked and numerical stability is evaluated with respect to penalty values. The following one-dimensional problem is considered:

$$\begin{cases} -\Delta u(x) = \left(4x^3 - 4x^2 - 6x + 2\right)\exp\left(-x^2\right), & \forall x \in (0,1) \\ u(0) = 1, \\ u(1) = 0. \end{cases} \tag{IV.1}$$

The analytical solution is $u_{\text{ex}}(x) = (1-x)\exp\left(-x^2\right)$. Numerical errors and convergence order in L^2-norm are computed for NIPG, SIPG and IIPG methods with linear to cubic polynomial approximations. Penalty values are such that $\sigma^{\text{I}} = \sigma^{\text{D}} = 1$ for SIPG with $p = 3$, NIPG and IIPG but $\sigma^{\text{I}} = \sigma^{\text{D}} = 2$ for SIPG with $p < 3$. Uniform meshes of size h are considered. The numerical error is defined as:

$$e_h = u_{\text{ex}} - u_h. \tag{IV.2}$$

The L^2-norm is:

$$\|e_h\|_{L^2(0,1)} = \left(\int_0^1 {e_h}^2(x)\,\mathrm{d}x\right)^{\frac{1}{2}}. \tag{IV.3}$$

Orders of convergence are obtained numerically by computing $\dfrac{\log(\|e_{h_1}\|/\|e_{h_2}\|)}{\log(h_1/h_2)}$. Theoretical convergence rates of primal DG method for uniform meshes can be derived. In one dimension , they are given for example in [26, 27], and reported in Tab. IV.1. Numerical errors and convergence rates are shown in Tab. IV.2.

Method	NIPG	SIPG	IIPG
Theoretical order	$p+1$ if p odd p if p even	$p+1$	$p+1$ if p odd p if p even

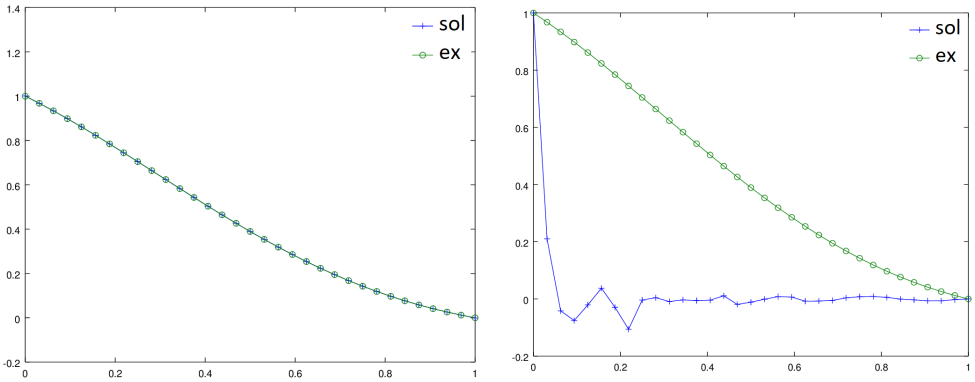
Table IV.1 – Theoretical convergence rates for 1D uniform meshes

		Linear approximation		Quadratic approximation		Cubic approximation	
Method	h	$\|e_h\|_{L^2(0,1)}$	Order	$\|e_h\|_{L^2(0,1)}$	Order	$\|e_h\|_{L^2(0,1)}$	Order
NIPG	1/2	9.7104E-02	-	1.2899E-02	-	6.2398E-04	-
	1/4	2.6019E-02	1.8999	4.1564E-03	1.6339	4.2126E-05	3.8887
	1/8	6.5796E-03	1.9835	1.2318E-03	1.7545	2.7853E-06	3.9188
	1/16	1.6361E-03	2.0078	3.2297E-04	1.9314	1.7792E-07	3.9685
	1/32	4.0652E-04	2.0088	8.1698E-05	1.9830	1.1217E-08	3.9875
SIPG	1/2	1.6443E-01	-	2.0337E-02	-	6.0659E-04	-
	1/4	3.4034E-02	2.2724	2.4644E-03	3.0448	2.3031E-05	4.7191
	1/8	7.0833E-03	2.2645	1.7036E-04	3.8545	1.1354E-06	4.3423
	1/16	1.3566E-03	2.3844	1.4595E-05	3.5450	5.3777E-08	4.4000
	1/32	2.4925E-04	2.4444	1.2907E-06	3.4992	2.4664E-09	4.4465
IIPG	1/2	1.37511E-01	-	4.4606E-02	-	4.4968E-03	-
	1/4	3.53440E-02	1.9600	7.3239E-03	2.6066	2.2775E-04	4.3034
	1/8	8.67816E-03	2.0260	1.5210E-03	2.2676	1.3408E-05	4.0863
	1/16	2.11935E-03	2.0338	3.5286E-04	2.1079	8.1252E-07	4.0445
	1/32	5.20962E-04	2.0244	8.5108E-05	2.0517	4.9747E-08	4.0297

Table IV.2 – Numerical errors and order of convergence for solution of problem Eq. (IV.1)

The numerical experiments are in good agreement with theoretical order and with those made in [26] on the same problem. This demonstrates the capability to reach high order approximation using DG methods. The code is written to work either from monomial basis or Lagrangian basis functions (classic finite element basis functions). Results do not differ in both case which is in accordance with DG methods because basis functions of $S_p(\mathcal{E}_h^n)$ have their support contained in one single element. This raises the question of the existence of a set of polynomial basis functions more suitable for a given problem or a specific use. A discussion about this topic can be found in [144]. In particular, modal basis like monomial basis are preferred to reach high order even though degrees of freedom loose their physical meaning. Orthogonality property is also often desired to diagonalize the mass matrix for instance.

Penalty values were varied during tests on the problem Eq. (IV.1). One observation is that there is a well-known threshold value below which instability happens, see [119] for details. Moreover, another behaviour was observed: some spurious effects occur for great penalty values beyond a threshold value. Numerical results for SIPG method with linear approximation on a grid with 32 elements are presented in Fig. IV.1 for two cases with different penalty values. As shown in Fig. IV.1b, the numerical solution becomes unstable. One possible explanation is that condition number of the system worsen the numerical scheme as stated in [83]. This aspect should be tackled in future investigations.



(a) Numerical solution for $\sigma^{\mathrm{I}} = \sigma^{\mathrm{D}} = 2$ (b) Numerical solution for $\sigma^{\mathrm{I}} = \sigma^{\mathrm{D}} = 10^{15}$

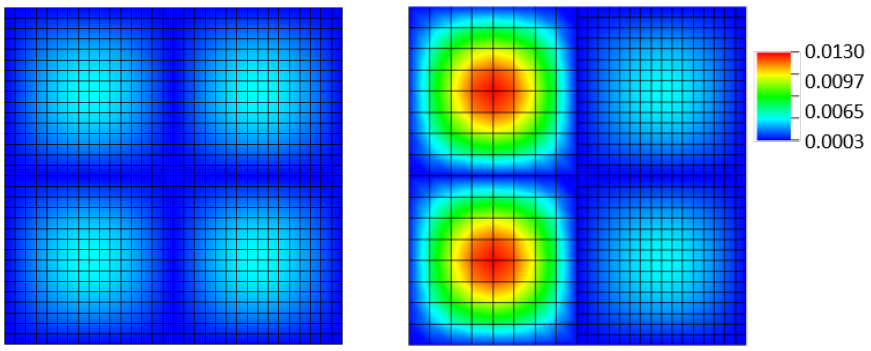
Figure IV.1 – Numerical error for model problem Eq. (IV.1) with SIPG method and $p = 1$.

IV.1.2 Two-dimensional analytical tests

Now, this second experiment focuses on the property of DG methods to treat non-conforming mesh. Only SIPG method will be treated seeing that there is no difference between the different methods in this case. High-order elements could be used but the experiment will be limited to linear approximation for brevity. This second experiment is considering the following two-dimensional problem, where $(x, y) \in \Omega = [0, 1] \times [0, 1]$:

$$\begin{cases} -\Delta u(x) = 8\pi^2 \sin(2\pi x) \sin(2\pi y), & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega. \end{cases} \tag{IV.4}$$

The analytical solution is $u_{\mathrm{ex}}(x, y) = \sin(2\pi x) \sin(2\pi y)$. Computations are done for SIPG method with $\sigma^{\mathrm{I}} = \sigma^{\mathrm{D}} = 20$. First of all, the numerical solution is computed on a conforming mesh of 1024 elements on which $\|e_h\|_{\mathcal{E}_h} = 1.6085 \times 10^{-3}$ (Fig. IV.2a). Then, a non-conforming mesh is used. The half part on the right is refined. Mesh is made of 696 elements and $\|e_h\|_{\mathcal{E}_h} = 3.2272 \times 10^{-3}$ (Fig. IV.2b). Numerical results are presented in Fig. IV.2.



(a) Conforming mesh (b) Non-conforming mesh

Figure IV.2 – Numerical error for model problem Eq. (IV.4) with SIPG method and $p = 1$.

Numerical comparison with the example given in [27] is in good agreement even though a quadrilateral mesh is used instead of triangles. An interesting point of view about the influence of mesh

nature can be found in Wirasaet *et al.* [191]. These results illustrate that DG methods can handle easily non-conforming mesh, giving path to mesh adaptation.

IV.2 Haverkamp's test-case

From now on, simulations are carried for Richards' equation given in the problem Eq. (II.1). A series of 1D infiltration cases is considered to validate the code and assess the method.

This numerical test-case was firstly introduced by Celia *et al.* [85]. It is based on an experiment from Haverkamp *et al.* [192] who mention the existence of quasi-analytical solution given by Philip [193]. It was taken afterwards by others like [19, 84, 94, 133, 194, 195] and corresponds to a class of classical test cases, for example, see [20]. Although the case is simple, it allows for understanding fundamental physics of a wetting front caused by infiltration. This is a first step to validate the code. In particular, the ability of the error indicator Eq. (III.2) to follow the front will be evaluated.

This problem deals with an 1D-infiltration into a soil column of 40 cm height and 8 cm width. On the top and bottom, Dirichlet boundary conditions for hydraulic head are prescribed: $h_{\text{top}} = 19.3$ cm, $h_{\text{bottom}} = -61.5$ cm which results in a cumulative downward infiltration. Sides are impervious. The initial condition is $h_0 = -61.5 + z$ cm. Simulation parameters are $p = 1$, $\sigma^{\text{I}} = \sigma^{\text{D}} = 100$ and $T = 600$ s. Threshold values for AMR selection from Section III.1.3 are: $\beta_{\text{c}} = \beta_{\text{r}} = 0.5$. The Haverkamp's test-case is sketched in Fig. IV.3.

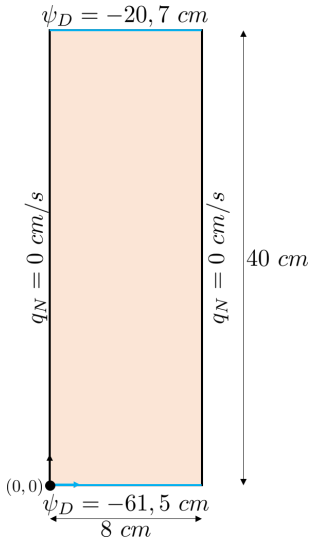
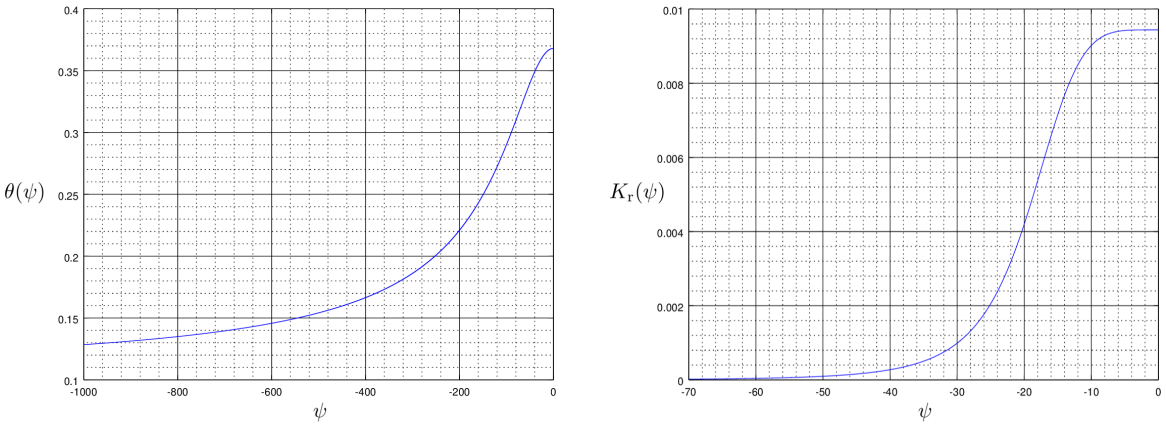


Figure IV.3 – Haverkamp's test-case configuration.

Hydraulic properties use Vachaud's relations (see Tab. I.1) with $A = 1.175 \times 10^6$, $B = 4.74$, $C = 1.611 \times 10^6$, $D = 3.96$, $K_{\text{s}} = 0.0094$ cm/s, $\theta_{\text{s}} = 0.287$ and $\theta_{\text{r}} = 0.075$. The shape of hydraulic properties functions given by this set of parameters is presented in Fig. IV.4 where one can see no steep gradient *a priori*.

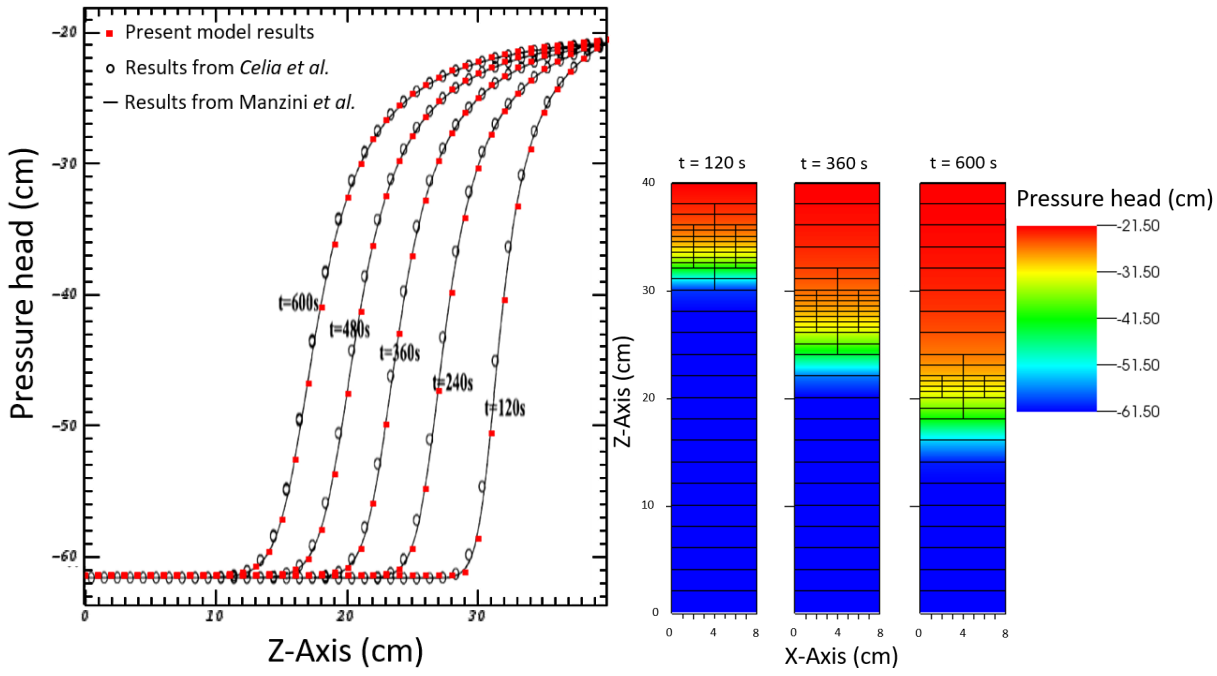


(a) Water content θ in function of pressure head ψ

(b) Relative hydraulic conductivity K_r in function of pressure head ψ

Figure IV.4 – Vachaud's relations for Haverkamp's test-case.

Two computations are performed. The first one uses a fixed mesh of 40 elements distributed vertically and the second one uses a mesh of 20 elements adding refinement based on the error indicator Eq. (III.2). The obtained numerical solution is in very good agreement to those from Celia *et al* and Manzini *et al.* [94] as seen in Fig. IV.5a. Moreover, mesh adaptation is able to capture the wetting front which is moving dynamically thanks to the error indicator Eq. (III.2) as it is shown in Fig. IV.5b.



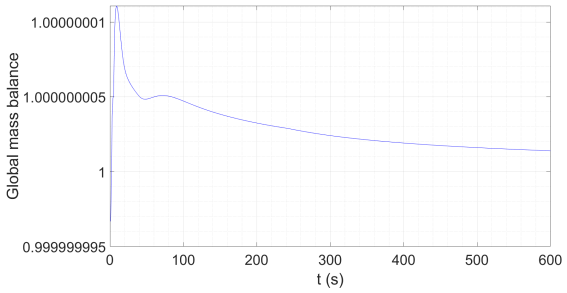
(a) Pressure head profiles for the fixed mesh of 40 elements

(b) Adaptive mesh refinement around the wetting front

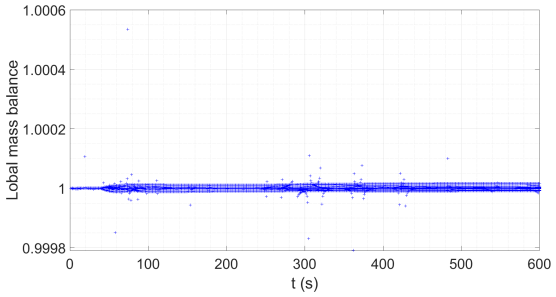
Figure IV.5 – Results for Haverkamp's test-case.

Local mass balance of Eq. (II.48) is evaluated for the Haverkamp's test-case. Global mass balance showed in Fig. IV.6a is respected almost perfectly. Element-wise mass balance is also very good during

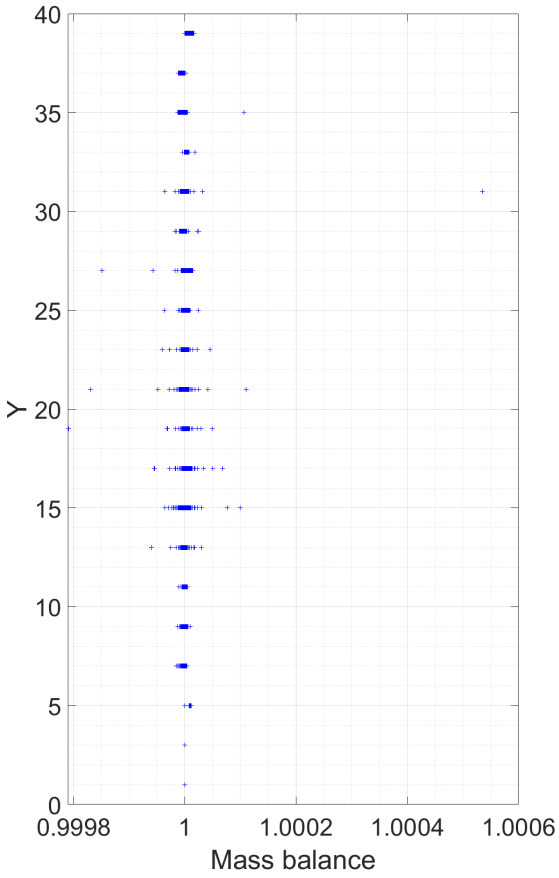
the whole simulation for most of elements. Figure IV.6b highlights local mass balance consistency over time and Fig. IV.6c shows every element along the soil column react identically.



(a) Global mass balance



(b) Local mass balance



(c) Distribution of local mass balance

Figure IV.6 – Mass balance for Haverkamp's test-case.

IV.3 Polmann's test-case

This test-case considers a soil from New Mexico whose hydraulic properties are provided by Polmann *et al.* [196]. A 1D vertical sample of this soil subject to downward infiltration was simulated by Celia *et al.* [85], Manzini and Ferraris [94] as well as Sochala [84]. This test-case can be challenging because, on one hand, the prescribed pressure head on both sides of the 100 cm soil column has a difference of 925 cm resulting in steep solicitation, and, on the other hand, hydraulic conductivity shows strong variations under the set of values taken by pressure head. The computational domain Ω is a rectangle $(0, 20) \times (0, 100)$ cm. The test-case is solved for pressure head ψ during $T = 172800$ s (48 h) with a constant time step $\tau = 120$ s and the initial condition is $\psi_0 = -1000$ cm. Implicit Euler scheme is used with $p = 1$, $\sigma_{\mathrm{I}} = \sigma_{\mathrm{D}} = 100$. Computation is done for two meshes: M100 is a coarse mesh of 100 elements and M1000 is a refined mesh of 1000 elements. The test-case configuration is sketched in Fig. IV.7.

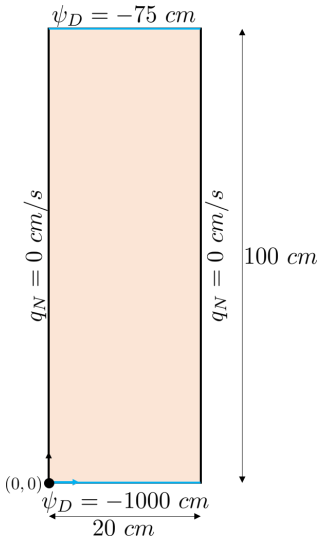
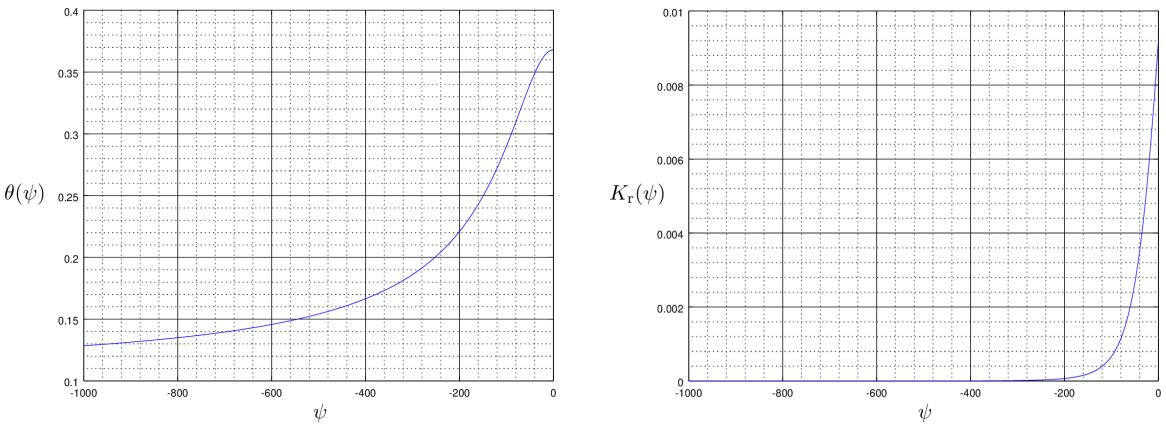


Figure IV.7 – Polmann's test-case configuration.

The Polmann's test-case employs Van Genuchten-Mualem relations (see Tab. I.1) with $K_s = 9.22 \times 10^{-3}$ cm·s^{-1}, $\theta_s = 0.368$, $\theta_r = 0.102$, $\alpha = 3,35 \times 10^{-2}$ cm^{-1}, $n = 2$, $m = 0.5$ and $l = 0.5$. Hydraulic properties functions are plotted in Fig. IV.8 where one can see a more brutal transition compared to hydraulic properties used in Haverkamp's test-case (Fig. IV.4).



(a) Water content θ in function of pressure head ψ

(b) Relative hydraulic conductivity K_r in function of pressure head ψ

Figure IV.8 – Van Genuchten-Mualem relations for Polmann's test-case.

For the mesh M100, nonlinear iterations have difficulties to reach the requirements of stopping criteria and the solution holds spurious oscillations, in particular an undershoot ahead of the sharp wetting front as showed in Fig. IV.9. This behaviour was already noticed by Celia *et al.* [85] and Sochala [84]. The latter decides to eliminate this undershoot by adding a slope limiter which causes a small late compared to the non-limited solution.

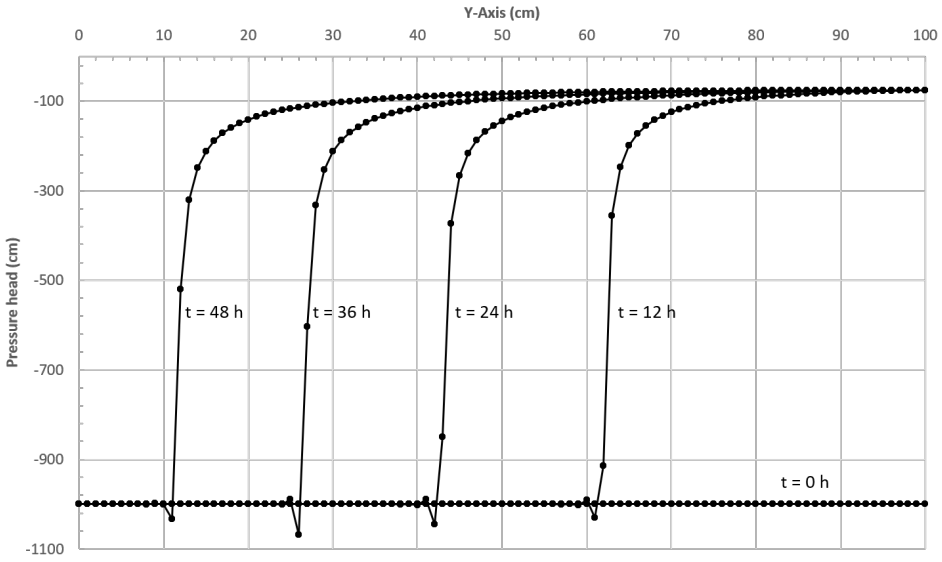


Figure IV.9 – Pressure head along the vertical for Polmann's test-case with the mesh M100.

Better results are obtained with the mesh M1000, the oscillations in the solution vanish, see Fig. IV.10. The M1000-results agree with those from Celia *et al.* [85] as well as from Manzini and Ferraris [94]. This shows that a good discretization is necessary for Richards' equation to get accurate solution. Such requirement can be fulfilled at optimal cost by a suitable adaptive local refinement. The error indicator from Eq. (III.2) is employed with refinement threshold values $\beta_c = \beta_r = 50$. Adaptation is performed every 5 time steps, starting with the M100 mesh. In Fig. IV.10, results show that refinement is able to capture the wetting front dynamically and eliminate the undershoot. The average number of elements is 211. More frequent adaptations allow to follow perfectly the front while less frequent adaptations are less costly.

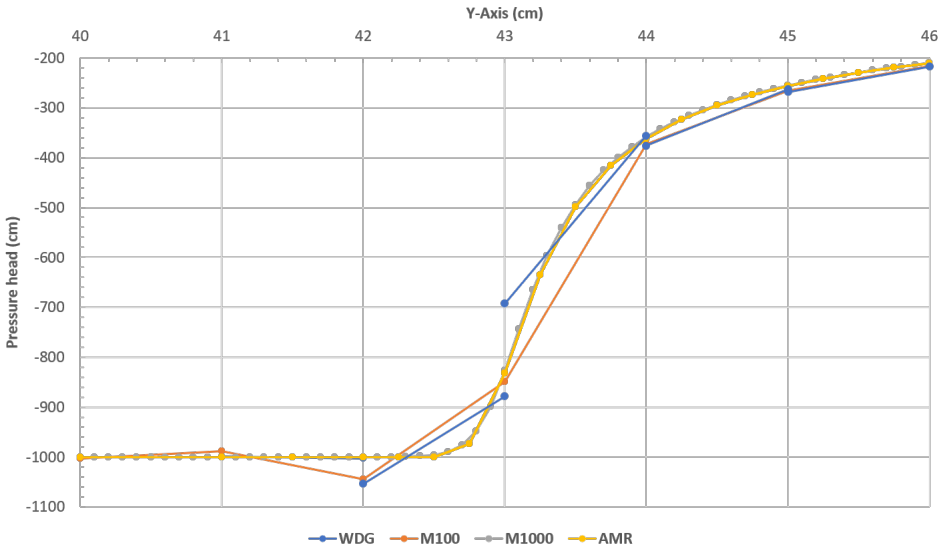


Figure IV.10 – Comparison of pressure head at $t = 24$ h for Polmann's test-case: IIPG on mesh with 100 elements (M100), IIPG on mesh with 1000 elements (M1000), IIPG with adaptive mesh refinement (AMR), WDG with 100 elements (WDG).

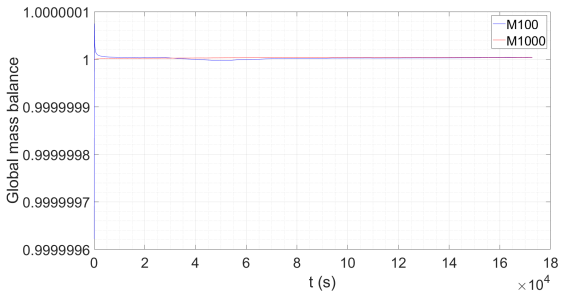
In addition, the WDG method from section III.3.1 is tried without any adaptation. The mesh M100 is used. Results in Fig. IV.10 show that the undershoot is still present but reduced to one single element. It is relevant to point out that the solution holds jumps where the wetting front is steep. They prevent oscillations to propagate and make solution catch the front steepness. The remaining undershoot depends on the wetting front localization in relation with the mesh geometry. This is due to the nonlinear nature of diffusivity and that is why it is important to associate mesh adaptation.

The extrapenalty term from Eq. (III.14) has no repercussion. Oscillations are not absorbed or reduced, no matter if it is employed solely or in conjunction with WDG method or mesh adaptation. However, convergence is affected: reaching stopping criteria is more difficult.

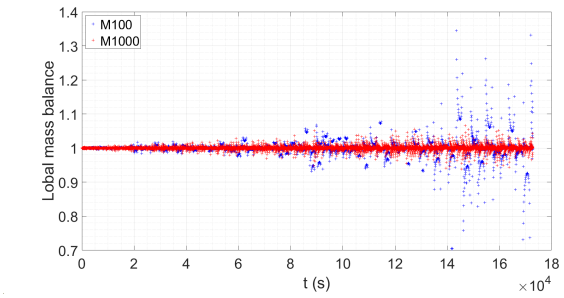
Local mass balance of Eq. (II.48) is evaluated for the Polmann's test-case with meshes M100 and M1000. Firstly, the global mass balance showed in Fig. IV.11a is respected almost perfectly in any cases with values very near to one. When mass balance is looked element-wise, some variations are observed both in space and time. As the simulation is running, local mass balance degrades over time as illustrated in Fig. IV.11b. Figure IV.11c shows that local mass balance is affected for elements at the bottom of the soil column. While local mass balance remains acceptable for mesh M1000, huge deviations appear for mesh M100. These discrepancies in local mass balance may be caused by the spurious oscillations of the sharp wetting front. Indeed, they follow the wetting front when it hits dry regions and are reduced with refinement. This indicates insufficient discretization for the scheme to respect mass balance in this particular case where high water content variations reach water-less regions. Roughly speaking, information should propagate from upwind to downwind flow but the downwind answer is not zero in the numerical scheme because of the sharpness of the front compared to the local mesh.

IV.4 Vogel-Van Genuchten-Cislerova test-cases

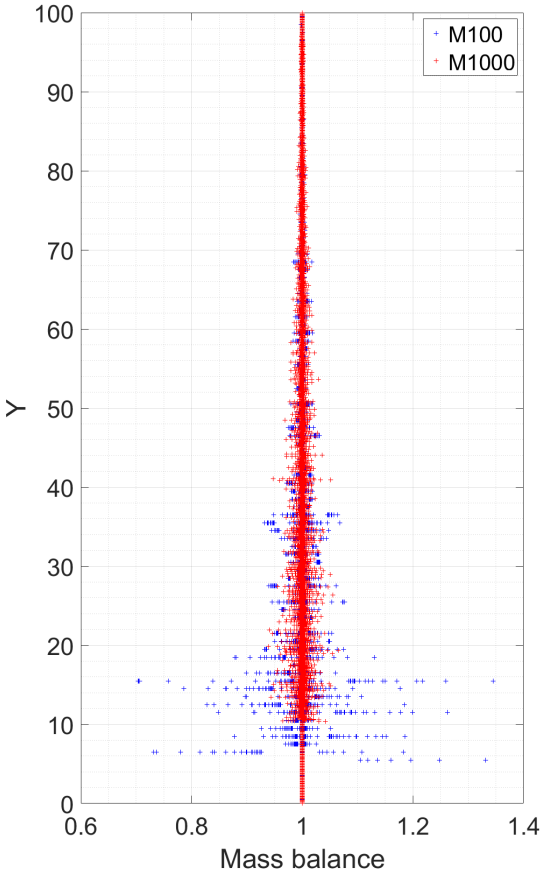
Vogel *et al.* [74] made three test-cases whose principle is the same: a 1D infiltration into a column of soil. After them, Sochala [84] used these tests to validate his code, Ginzburg *et al.* [104] assess a lattice Boltzmann method for Richards' equation on one of them and De Maet *et al.* [124] demonstrate their numerical method with a similar case from the code HYDRUS [136]. Even if these test-cases are one-dimensional infiltration, they are interesting for two main reasons apart from being quite classical. One is because they employ other types of boundary conditions like a non-zero Neumann boundary condition or Dirichlet boundary conditions resulting in upward infiltration. The second reason is because they employ hydraulic properties based on non-standard relations, called modified Van Genuchten-Mualem relations (see Tab. I.1), including a parameter capable of affecting their shape. Vogel *et al.* [74] introduce this so-called minimum capillary height ψ_s to provide more flexible relations for hydraulic properties and model better the capillary zone behaviour, when the porous medium is almost fully saturated but with negative pressure head. This parameter acts as the air entry pressure ψ_e but it can take smaller values, about centimetres, which corresponds to maximum pore size. Use and validity of these relations are not discussed here. The goal is to evaluate the model answer to such nonlinear relations. If $\psi_s = 0$, the modified Van Genuchten-Mualem relations are equivalent to the standard Van Genuchten-Mualem relations.



(a) Global mass balance



(b) Local mass balance

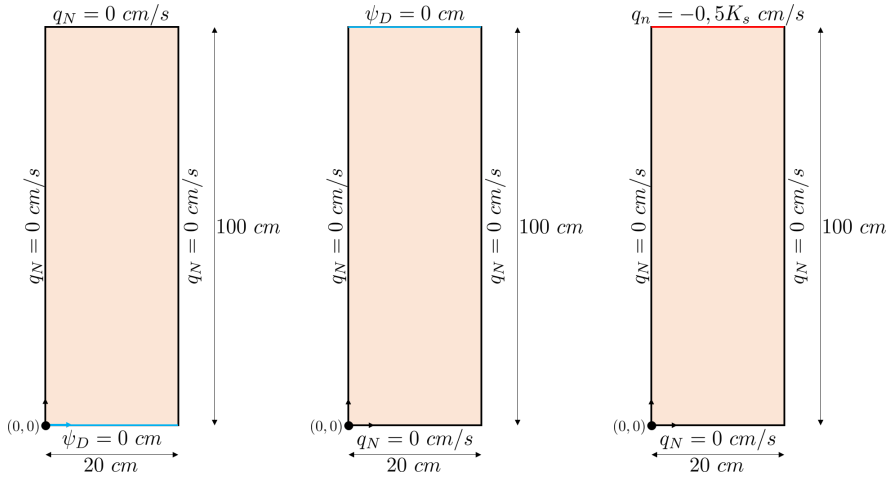


(c) Distribution of local mass balance

Figure IV.11 – Mass balance for Polmann's test-case. Blue values are for the mesh M100 and red values for the mesh M1000.

Geometry is the same for the three cases with Ω being a rectangle $(0, 20) \times (0, 100)$ cm. Simulations use $\tau = 120$ s, $p = 1$, $\sigma^{\mathrm{I}} = \sigma^{\mathrm{D}} = 100$ and 100 elements distributed vertically. The initial condition is hydrostatic: $h_0 = -1000$ cm. Boundary conditions are different which results in different infiltrations defining the three test-cases (see Fig. IV.12):

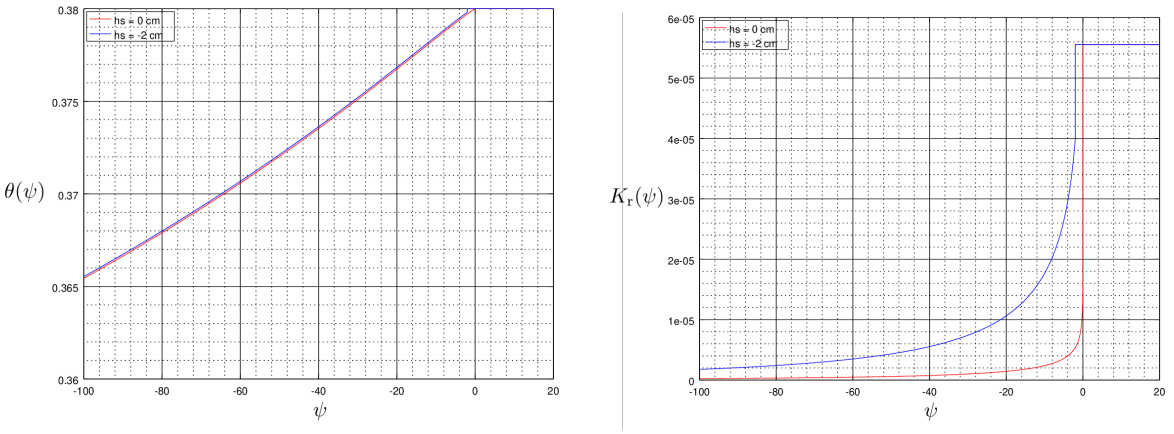
- First test-case: upward infiltration. The simulation lasts $T = 86400$ s $= 24$ h. Pressure head is plotted every 17280 s $= 0.2$ d;
- Second test-case: ponded downward infiltration. The simulation lasts $T = 8640$ s $= 0,1$ d. Pressure head is plotted every 43200 s$= 12$ h;
- Third test-case: flux-controlled downward infiltration. The simulation lasts $T = 8640$ s $= 0,1$ d. Pressure head is plotted every 43200 s$= 12$ h.



(a) Upward infiltration (b) Ponded downward infiltration (c) Flux-controlled downward infiltration

Figure IV.12 – Vogel-Van Genuchten-Cislerova test-cases configuration.

Moreover, each test-case is run for both $\psi_s = 0$ cm and $\psi_s = -2$ cm. Other parameters for hydraulic properties are $K_s = 5.5556 \times 10^{-5}$ cm/s, $\theta_s = 0,38$, $\theta_r = 0,068$, $\alpha = 8 \times 10^{-3}$ cm^{-1}, $n = 1,09$ and $l = 0.5$. The functions are represented in Fig. IV.13. Hydraulic conductivity is more sensible than water content for a variation of minimum capillary height ψ_s.



(a) Water content θ in function of pressure head ψ (b) Relative hydraulic conductivity K_r in function of pressure head ψ

Figure IV.13 – Modified Van Genuchten-Mualem relations for Vogel-Van Genuchten-Cislerova test-cases.

The numerical results presented in Fig. IV.14 are in good agreement with those of Vogel *et al.* [74], Ginzburg *et al.* [104] and Sochala [84]. These three test-cases are interesting to evaluate the performance of nonlinear solving for the DG scheme applied to Richards' equation. Indeed, for $\psi_s = 0$ cm, hydraulic conductivity holds strong variations when the porous medium is near saturation, illustrating what a capillary-barrier effect is. For $\psi_s = -2$ cm, the wetting front propagation is fast due to a higher hydraulic conductivity. It has been noted that the stopping criteria may have different order of magnitude between the cases which may trouble convergence if the user gives arbitrary values. The

cases with $\psi_s = 0$ cm are more demanding since hydraulic conductivity shape is steep. So, nonlinear iterations struggle to converge and adaptive time stepping is then suitable to reach convergence. It is particularly the case for the flux-controlled downward infiltration with $\psi_s = 0$ cm. Besides, the model manifests some front instabilities for downward infiltrations with $\psi_s = 0$ cm which confirms that something is needed when front is steep with insufficient discretization. No oscillation occurs for upward infiltration because the flux is against gravity taking values outside of the most nonlinear portion of hydraulic conductivity.

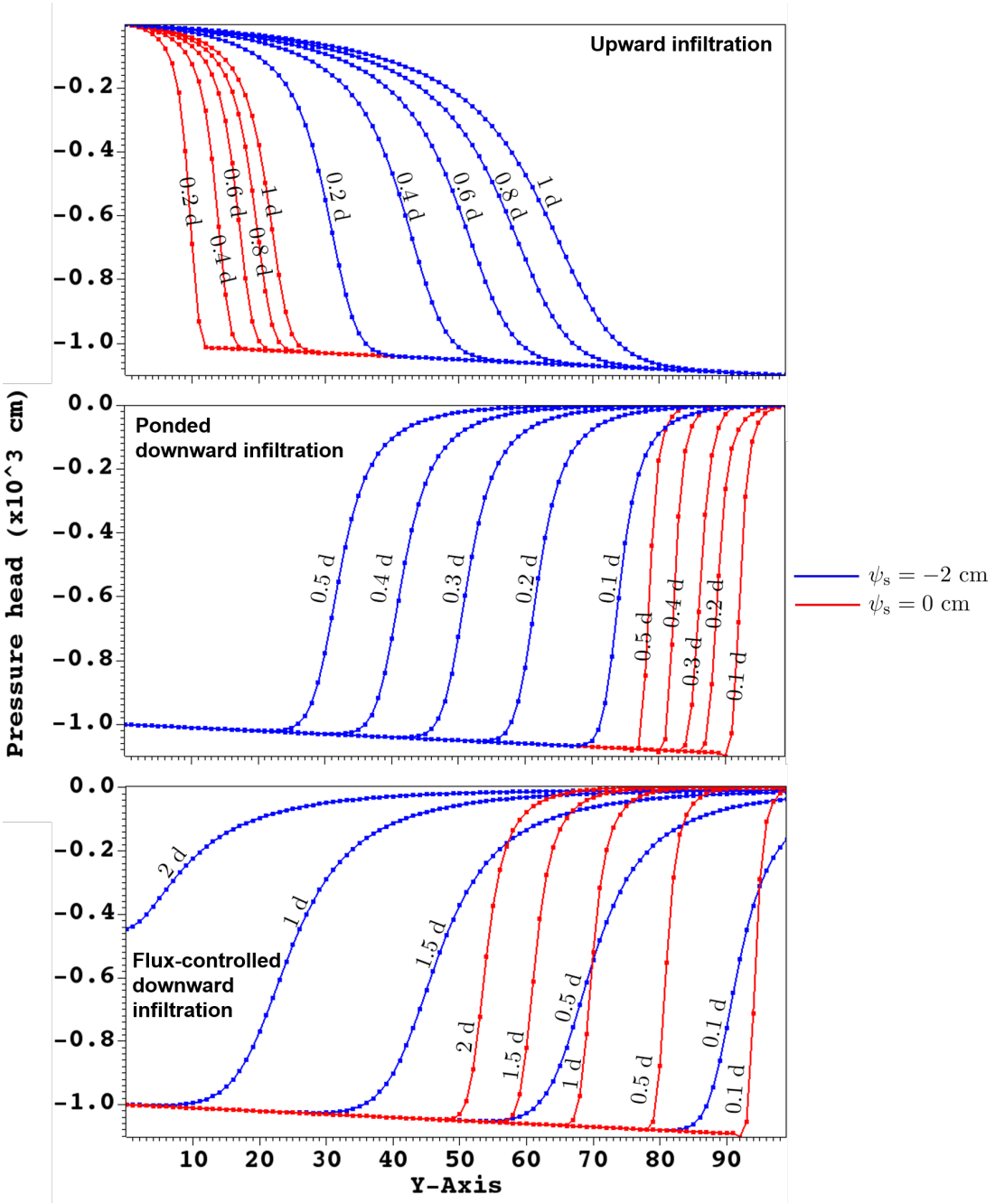


Figure IV.14 – Results for Vogel-Van Genuchten-Cislerova test-cases for $\psi_s = -2$ cm (blue) and $\psi_s = 0$ cm 'red) with upward infiltration (top), ponded downward infiltration (middle), flux-controlled downward infiltration (bottom).

IV.5 Tracy's benchmark

This test-case has an analytical solution given by Tracy [197] for 2D and 3D problems which can be used as a benchmark for Richards' equation [198, 199]. Tracy's benchmark is particularly relevant because it is transient, has a simple parameter to vary nonlinearities for Richards' equation and holds differentiated steep regions. It can serve to check the nonlinear solver robustness, convergence properties and AMR. Then, it is very convenient to assess the solution quality and compare codes: Šolín and Kuraz [112] and Dolejší *et al.* [108] used it to evaluate the performance of their respective method.

Tracy's benchmark employs Gardner-Irmay relations with $K_s = 1.1\ \mathrm{m{\cdot}d^{-1}}$, $\theta_\mathrm{s} = 0.5$, $\theta_\mathrm{r} = 0$, $\alpha = 0.1\ \mathrm{m}^{-1}$ and $m = 1$. It is solved for pressure head ψ. The residual pressure head ψ_r is a parameter. The computational domain Ω is a square $(0, a) \times (0, b)$ m. Here, $a = b = 1$ m and $\psi_\mathrm{r} = -10$ m. A specific Dirichlet boundary condition is prescribed on the top edge with the function:

$$\psi_\mathrm{top} = \frac{1}{\alpha} \log\left(e^{\alpha\psi_\mathrm{r}} + \left(1 - e^{\alpha\psi_\mathrm{r}}\right) \sin\left(\frac{\pi x}{a}\right)\right). \tag{IV.5}$$

For other edges, a constant Dirichlet boundary condition is imposed: $\psi = \psi_\mathrm{r}$. The initial condition is $\psi_0 = \psi_\mathrm{r}$. Tracy gives the 2D exact solution ψ_{ex} [197]:

$$\psi_\mathrm{ex} = \frac{1}{\alpha} \log\left(e^{\alpha\psi_\mathrm{r}} + \psi_\mathrm{ts} + \psi_\mathrm{ss}\right), \tag{IV.6}$$

$$\psi_\mathrm{ss} = \frac{1}{\alpha} \log\left(e^{\alpha\psi_\mathrm{r}} + \left(1 - e^{\alpha\psi_\mathrm{r}}\right) \sin\left(\frac{\pi x}{a}\right) e^{\frac{\alpha}{2}(b-y)} \frac{\sinh \beta_\mathrm{T} y}{\sinh \beta_\mathrm{T} b}\right), \tag{IV.7}$$

$$\psi_\mathrm{ts} = \frac{2\left(1 - e^{\alpha\psi_\mathrm{r}}\right)}{b\gamma_\mathrm{T}} \sin\left(\frac{\pi x}{a}\right) e^{\frac{\alpha}{2}(b-y)} \sum_{k=1}^{\infty} (-1)^k \frac{\lambda_k}{\mu_k} \sin(\lambda_k y) e^{-\mu_k t}, \tag{IV.8}$$

$$\beta_\mathrm{T} = \sqrt{\left(\frac{\alpha}{2}\right)^2 + \left(\frac{\pi}{a}\right)^2}, \tag{IV.9}$$

$$\gamma_\mathrm{T} = \frac{\alpha(\theta_\mathrm{s} - \theta_\mathrm{r})}{b}, \tag{IV.10}$$

$$\lambda_k = \frac{k\pi}{b}, \tag{IV.11}$$

$$\mu_k = \frac{\beta_\mathrm{T}^2 + {\lambda_k}^2}{\gamma}, \tag{IV.12}$$

where ψ_ss denotes the steady-state solution and ψ_ts the transient counterpart for the exact solution. The configuration is sketched in Fig. IV.15. Figure IV.16 shows the numerical solution obtained for a fixed grid with $\tau = 10^{-4}$ day and $T = 10^{-2}$ day.

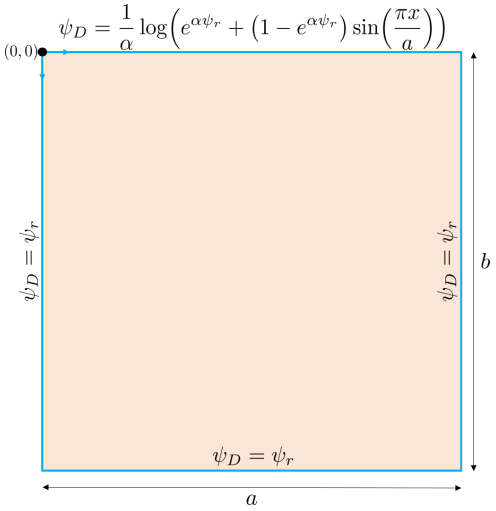


Figure IV.15 – Tracy's benchmark configuration.

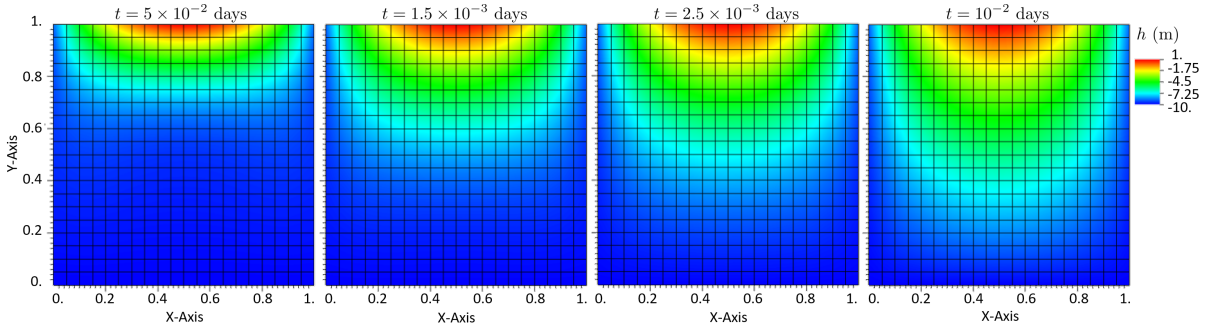


Figure IV.16 – Hydraulic head for Tracy's benchmark with $p = 1$ and 6-step BDF.

The treatment of initial condition is not straightforward. First, because the initial condition does not respect the top boundary condition: they are inconsistent. This issue can be troublesome for discretizations like finite elements methods [112] but it is natural for DG methods where boundary conditions are usually enforced by weak penalization [108]. Second, because the top boundary condition is particularly stiff compared to the initial state. Therefore, the very first moments are the most interesting to simulate in order to address the time-boundary inconsistency and the early stiffness. Nonlinearities are controlled by the parameter α in Eq. (IV.5). The larger α, the faster the relative hydraulic conductivity drops to nearly zero, and so, nonlinearities are increased for Richards' equation [198]. Large values of α mean greater errors in the solution unless to reduce the time step [199]. It challenges the nonlinear solver so much that the Newton-Raphson method struggles to converge [199]. Moreover, errors are linked to the top corners since the top boundary condition induces a steep slope at these corners. This is particularly the case for smaller values of α and transient state but errors can also lie at the bottom of the grid for largest values of α and steady state [198, 199]. As Tracy stated, these features make this problem a good option to investigate AMR.

The classic L^2-norm and the energy norm, also called the DG intrinsic norm, are introduced:

$$\|u\|^2_{L^2(E)} := \int_E u^2 \,\mathrm{d}E \qquad |||u|||^2_{\mathcal{E}(E)} := \|u\|^2_{\mathrm{R}(E)} + \sum_{F\in\partial E} \|u\|^2_{\mathrm{J}(F)}, \tag{IV.13}$$

where the residual seminorm and the jump seminorm are respectively:

$$\|u\|_{\mathrm{R}(E)}^2 := \int_E \mathbb{K}(u)\nabla u \cdot \nabla u \,\mathrm{d}F = \left\|(\mathbb{K}(u))^{\frac{1}{2}}\nabla u\right\|_{L^2(E)}^2, \tag{IV.14}$$

$$\|u\|_{\mathrm{J}(F)}^2 := \int_F \varrho_F [\![u]\!]^2 \,\mathrm{d}F = \varrho_F \|[\![u]\!]\|_{L^2(F)}^2. \tag{IV.15}$$

Errors between numerical and exact solution are computed for a time step such that:

$$\|e\|_X = \|\psi_{\mathrm{ex}} - \psi\|_X \quad \text{with } X = \left\{L^2, \mathcal{E}, R, J\right\}. \tag{IV.16}$$

To measure how effective the estimation-based error indicator is, the effectivity indices for the estimation-based error indicator (III.4), the element and face residuals (III.5, III.6), and the solution jump (III.7) will be used:

$$I_{\mathcal{E}}^{\mathrm{eff}} := \frac{\eta_E^n}{|||e|||_{\mathcal{E}(E)}}, \qquad I_R^{\mathrm{eff}} := \frac{\left(\eta_{\mathrm{R}E}^{n\,2} + \eta_{\mathrm{F}E}^{n\,2}\right)^{\frac{1}{2}}}{\|e\|_{R(E)}}, \qquad I_J^{\mathrm{eff}} := \frac{\eta_{\mathrm{J}E}^n}{\|e\|_{J(F)}}. \tag{IV.17}$$

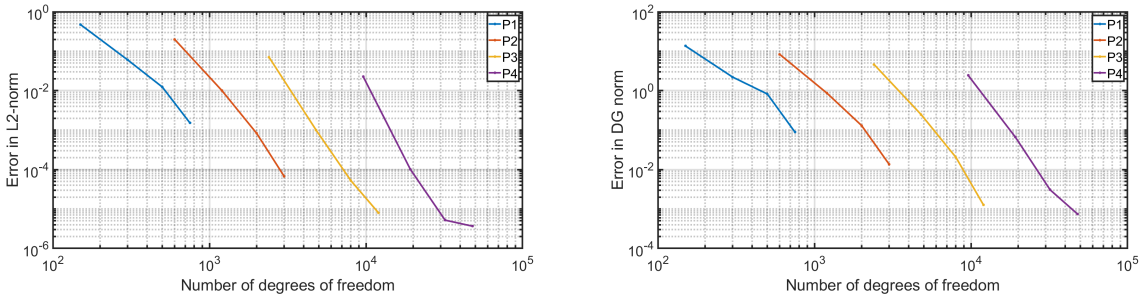
Due to the lack of rigorous derivation, the effectivity index is not expected to represent properly the true error. The estimator efficiency is rather appreciated to set up a mesh refinement strategy.

First of all, several simulations are performed to show convergence properties. Because of symmetry, only one half vertical plan of the domain is considered to speed up the simulation. Duration is set to $T = 10^{-4}$ day to focus on small times. The computations are carried out with the 6-step BDF and time steps are set to $\tau = 10^{-6}$ day. This is needed to reduce time discretization errors expecting that they will be negligible compared to space discretization errors in order to observe space convergence. Indeed, according to theoretical error estimates [27], error in L^2-norm and H^1-seminorm behave such that:

$$\|e\|_{L^2} \approx c_h h^p + c_\tau \tau^q, \tag{IV.18}$$

$$\|e\|_{\mathcal{E}} \approx c'_h h^{p+1} + c'_\tau \tau^q, \tag{IV.19}$$

where c_h, c'_h, c_τ and c'_τ are constants independent of h and τ. Numerical experiments involve four meshes of 50, 200, 400 and 3200 quadrilateral elements. Order approximation varies from one to four. Figure IV.17 shows the convergence of the error e_h for the L^2- and energy norms with respect to the number of degrees of freedom. It is interesting to note that space error are so small for the most dense mesh with $p = 4$ that time error saturates the convergence.

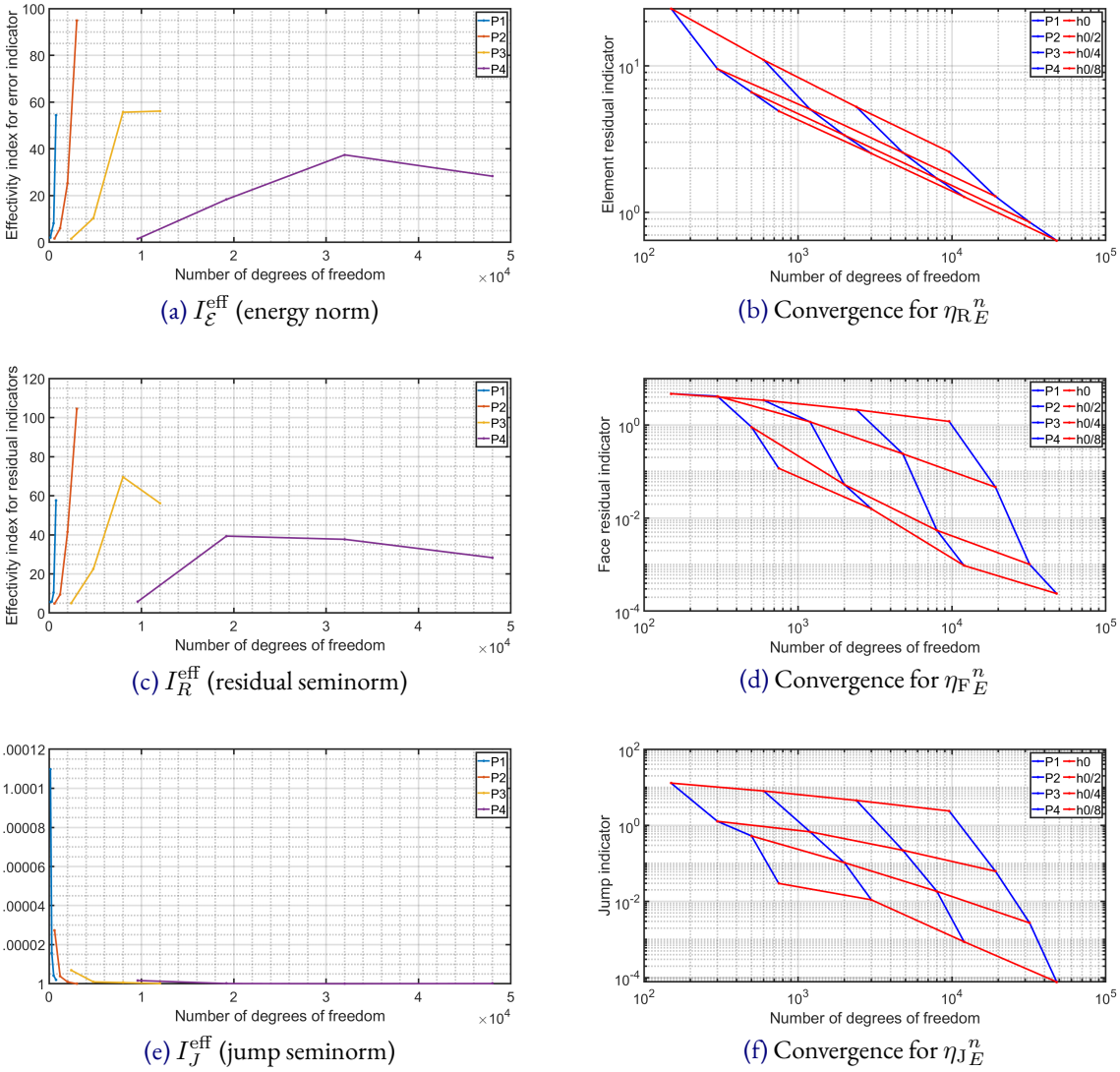


(a) Errors in the L^2-norm

(b) Errors in the energy norm

Figure IV.17 – Numerical convergence for Tracy's benchmark.

In Fig. IV.18, effectivity indices are plotted. While the jump indicator Eq. (III.7) is very effective, the other estimation-based error indicator components are not representative of the true error in term of magnitude. Nevertheless, they always overestimate the true error which is good for mesh refinement monitoring because the true error might be less than expected. The behaviour of the three contributions with respect to the number degrees of freedom is shown in Fig. IV.18 by varying the order of approximation and the mesh size. Since they do not represent properly the true error, observations are difficult to relate with the true error. However, some remarks can be drawn. Firstly, these diagrams show the convergence up to a certain extent. Secondly, one can see that the volume contribution – the residual indicator from Eq. (III.5) – behaves quite differently compared to the surface contributions – the flux and jump indicators from Eq. (III.5) and (III.6). The convergence is globally maximised with *p*-refinement for the surface contributions while *h*-refinement seems to improve convergence better for the volume contribution. This observation should be assessed in more details but could drive the *hp*-decision making.



(a) $I_{\mathcal{E}}^{\text{eff}}$ (energy norm)

(b) Convergence for $\eta_{\text{R}}{}_E^n$

(c) I_R^{eff} (residual seminorm)

(d) Convergence for $\eta_{\text{F}}{}_E^n$

(e) I_J^{eff} (jump seminorm)

(f) Convergence for $\eta_{\text{J}}{}_E^n$

Figure IV.18 – Effectivity indices for different indicators (left) and convergence diagrams for the three parts of the indicator η_E^n (right).

Error estimation is evaluated in terms of ability to drive the mesh adaptation. The error in L^2-norm is above the error measured with energy norm as stated in Fig. IV.19. It is worth to notice that the estimation-based error indicator is able to give the shape of the true error for the energy norm, particularly for steady state, see Fig. IV.20. Besides, the estimation-based error indicator spots the error in the top corner where Tracy stated there were problems for this test-case. The gradient-based error indicator was used successfully to capture the moving wetting front on the simple 1D infiltration Haverkamp's test-case, see Section IV.2. Yet, as shown in Fig. IV.21 and IV.22, it performs badly for the Tracy's benchmark where steep gradients arise on a large part of the domain. The refinement thresholds are difficult to tune which leads to over-refinement for large regions at the top. They were set up to $\beta_c = \beta_r = 0.5$ in order to give $|||e|||_{L^2(\Omega)} \approx 0.167$ at $t = 10^{-2}$ d. Then, the computation is very costly since it takes around 227 min. For the estimation-based error indicator, the ability to drive the mesh adaptation is more convincing. $\beta_c = \beta_r = 0.85$ in order to target $|||e|||_{L^2(\Omega)} = 0.137$ at $t = 10^{-2}$ d. The mesh is adapted dynamically without inducing over-refinement. Elements are refined where it is needed because adaptation occurs locally while the error is controlled. Then, the cost is reasonable with a computation around 35 min.

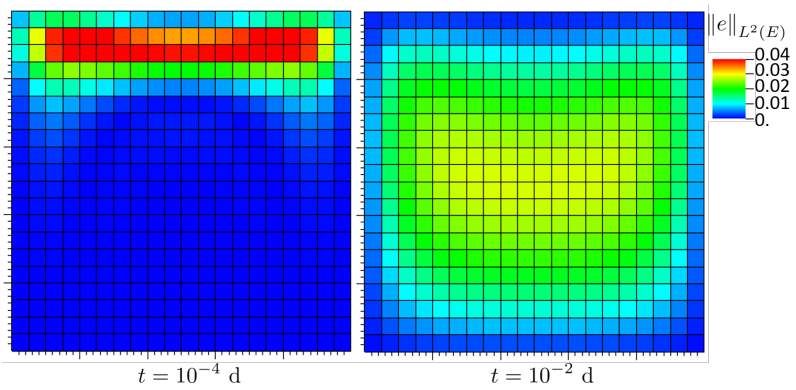
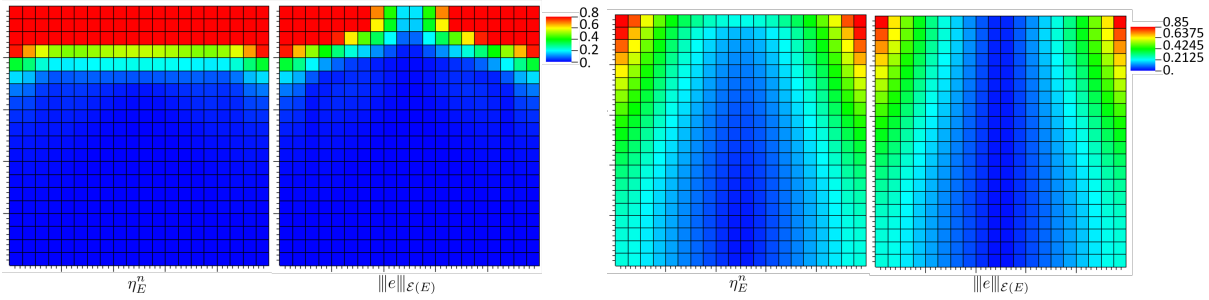


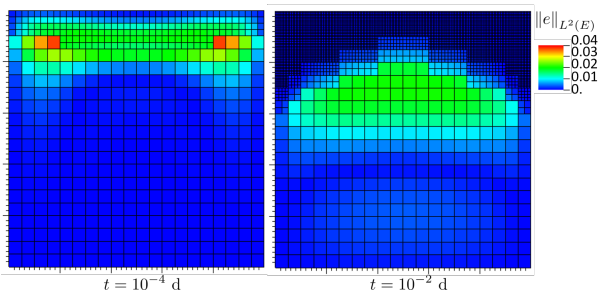
Figure IV.19 – True error distribution in the L^2-norm on Tracy's benchmark.



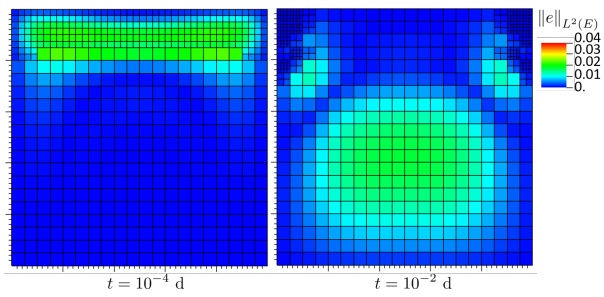
(a) Distribution at $t = 10^{-4}$ d

(b) Distribution at $t = 10^{-2}$ d

Figure IV.20 – Comparison between distributions of true error in energy norm and the indicator η_E^n on Tracy's benchmark.

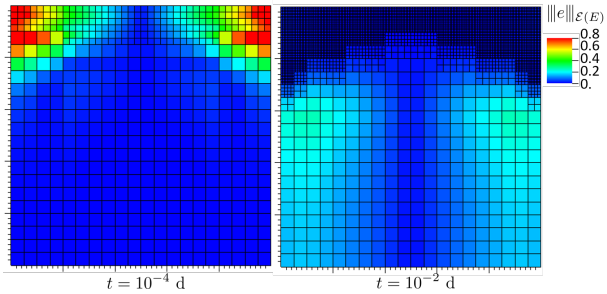


(a) Mesh and error distribution in L^2-norm for C_E-driven AMR

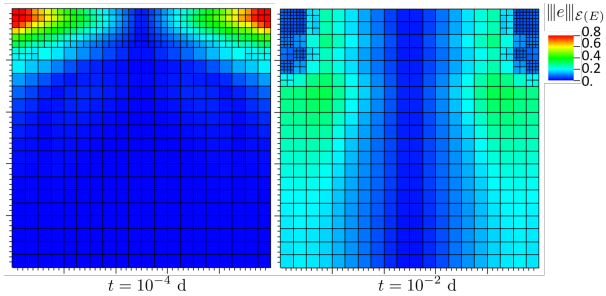


(b) Mesh and error distribution in L^2-norm for η_E^n-driven AMR

Figure IV.21 – Comparisonin with L^2-norm between C_E-based adaptation (left) and η_E^n-based adaptation on Tracy's benchmark.



(a) Mesh and error distribution in energy norm for C_E-driven AMR



(b) Mesh and error distribution in energy norm for η_E^n-driven AMR

Figure IV.22 – Comparison with energy norm between C_E-based adaptation (left) and η_E^n-based adaptation on Tracy's benchmark.

IV.6 Comparison with Vauclin's studies

Vauclin, Vachaud and Khanji made a series of laboratory experiments in the 1970's whose description can be found in [200–203]. They imply water table recharge and drainage in a slab of sandy soil.

IV.6.1 Water table recharge

Water flow recharge through a slab of soil is simulated in Vauclin *et al.* [203] in which experimental details and results can be found. It consists in a box of 6 m by 2 m whose only one half is simulated because of symmetry. Sides are impervious on the left, at the top for $x > 50$ cm and at the bottom: $\nabla h \cdot \boldsymbol{n} = 0$. A constant flux is prescribed on the top for $x \leq 50$ cm: $-\mathbb{K}(\psi)\nabla h \cdot \boldsymbol{n} = 14.8$ cm/h. The water level is kept constant in the ditch on the right for $z \leq 65$ cm: $h = 65$ cm. The remaining boundary on the right for $z > 65$ cm accounts for a seepage boundary condition. Initial state is at hydrostatic equilibrium with water table at $z = 65$ m.

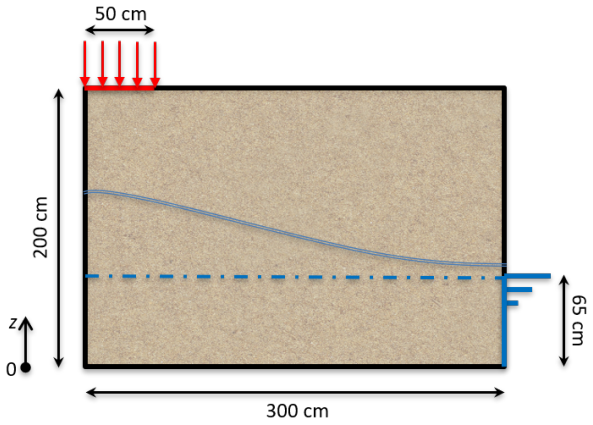


Figure IV.23 – Water table recharge test-case configuration.

Test-case configuration is depicted in Fig. IV.23. Hydraulic properties are taken from Vachaud's relations (see Tab. I.1) with $A = 2.99 \times 10^6$, $B = 5.0$, $C = 40000$, $D = 2.90$, $K_\mathrm{s} = 35$ cm/h, $\theta_\mathrm{s} = 0.30$ and $\theta_\mathrm{r} = 0$. This experiment will be simulated in two steps: a first simulation will catch the steady state and the other one takes the whole experiment with the transient state.

Steady state

The following steady-state Richards' equation is considered:

$$-\nabla \cdot (\mathbb{K}(h - z)\nabla(h + z)) = 0. \tag{IV.20}$$

Numerical simulation is performed with Lagrangian basis functions, linear approximation $p = 1$, $\sigma^\mathrm{I} = \sigma^\mathrm{D} = 20$ on a mesh of 2367 elements. Zones around the expected position of water table and boundary recharge flux are refined, see Fig. IV.24. Water table location matches $\psi = 0$ and capillary fringe location is established at $\theta(\psi) = 0.05\theta_\mathrm{s}$. Numerical results are shown in Fig. IV.25. Values and red contour lines of hydraulic head h together with distribution of flux are displayed.

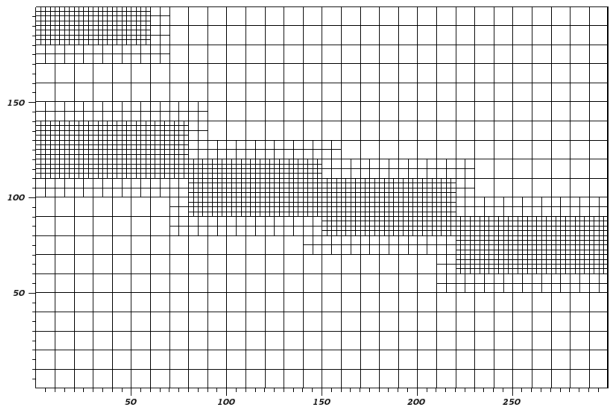


Figure IV.24 – Mesh used for the steady-state simulation of water table recharge test-case.

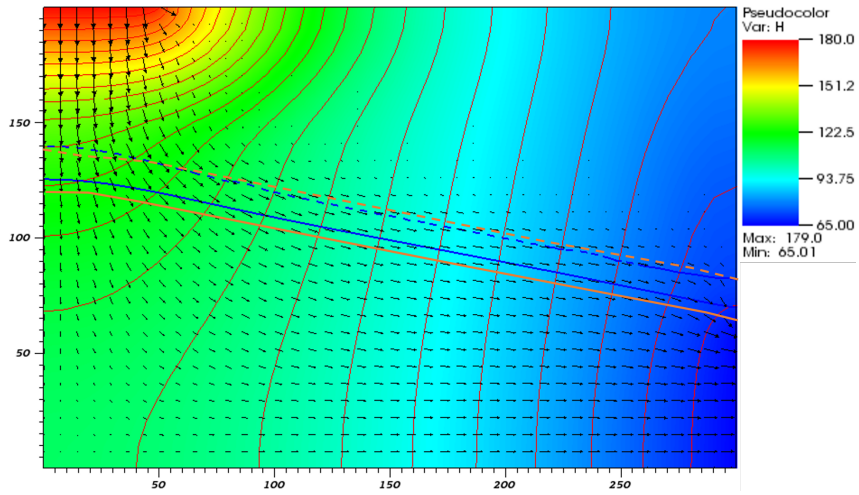


Figure IV.25 – Numerical simulation with distributions flux and hydraulic head h. Calculated water table and capillary fringe are respectively represented by a blue solid line and a blue dotted line. Those of Vauclin *et al.* [203] are represented by an orange solid line and an orange dotted line.

Water table, initially at $z = 65$ cm, raises because of the recharge which moistens the unsaturated zone. Compared to experiment [203], distribution of flux as well as positions of water table and capillary fringe are in good agreement even though a small shift is to notice. Distribution of hydraulic head is consistent with experiment within the saturated zone and under the recharge zone. However, on the top left-hand area, hydraulic head is higher than one measured and obtained by Vauclin *et al.* [203].

Transient state

The complete simulation of water table recharge by Vauclin et al. [203] was used by many other studies to assess their method, see [93, 204–208]. The validation of the code MODFLOW is partly done with this experiment dataset [19]. One can notice that the majority of these simulations does not take into account the seepage boundary condition and prescribe an impervious wall instead which is in contradiction with the experimental set-up and numerical studies of Vauclin *et al*. They argue that this problem a steady-state solution observed experimentally and numerically without the development of a significant seepage face [204].

The simulation presented now is carried for Richards' equation given in the problem Eq. (II.1). Slightly different parameters are changed compared to the steady state simulation: the mesh is a regular grid $\Delta x = \Delta z = 5$ cm, monomial basis are used with quadratic approximation and $\sigma^{\mathrm{I}} = \sigma^{\mathrm{D}} = 100$. Simulation lasts $T = 8$ h and the implicit Euler scheme with adaptive time stepping is used. The transient position of the water table is reported every 2 hours, as shown in Fig. IV.26, and compared with experimental results taken from [203]. The results obtained from the numerical model closely agree with the experimental profile. The growth of water table mound is well simulated. Small discrepancies are observed in the middle of water table which is underestimated by numerical results. This late which is observed by other study using different models, see [203, 206] and may be explained by the sandy soil which is non-perfectly isotropic and homogeneous as stated by hydraulic properties. Indeed, inflow from recharge flux is redistributed horizontally in the unsaturated zone by diffusion before reaching water table. Then, the unsaturated zone stores a portion of water inflow moistening dry regions and lateral water redistribution in the saturated zone is less significant. Similar observations are made in [205, 207]. One can see that seepage face length is very small and so, may be disregarded for this experiment as stated in [204].

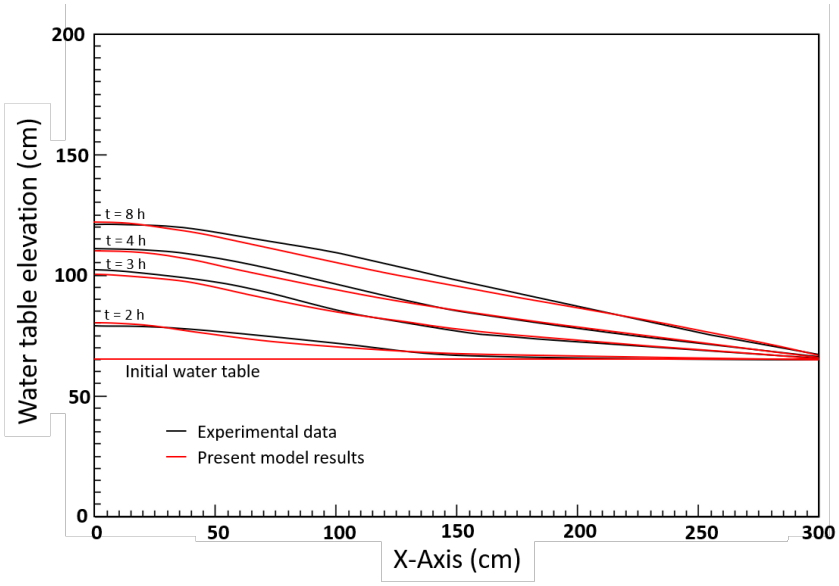


Figure IV.26 – Numerical simulation of two-dimensional recharge with water table position compared to experimental data from Vauclin *et al.* [203].

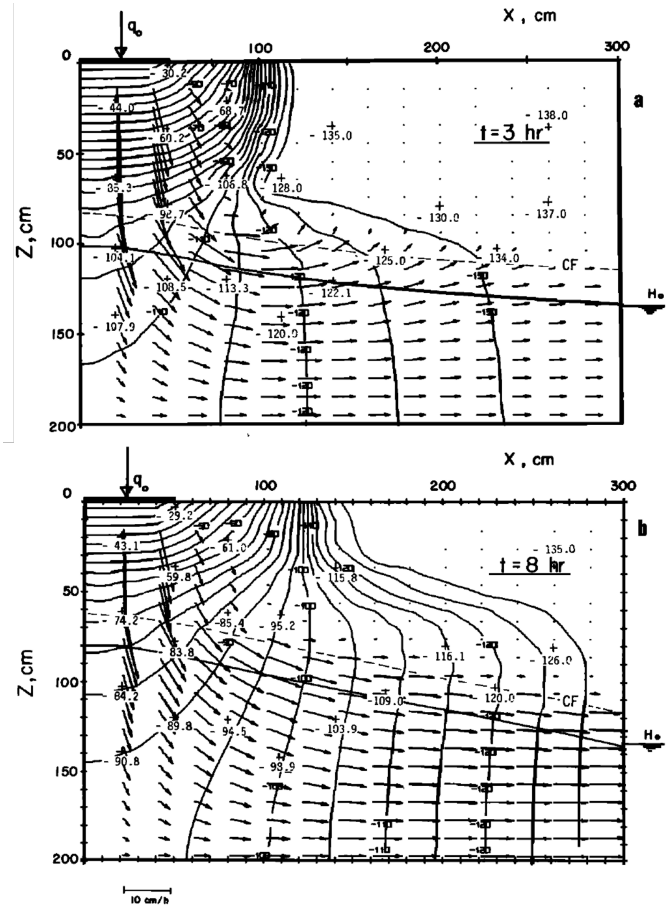


Figure IV.27 – Numerical distribution of hydraulic head (experimental data are numbers with crosses) and flux expressed in centimetres of water. After Vauclin *et al.* [203].

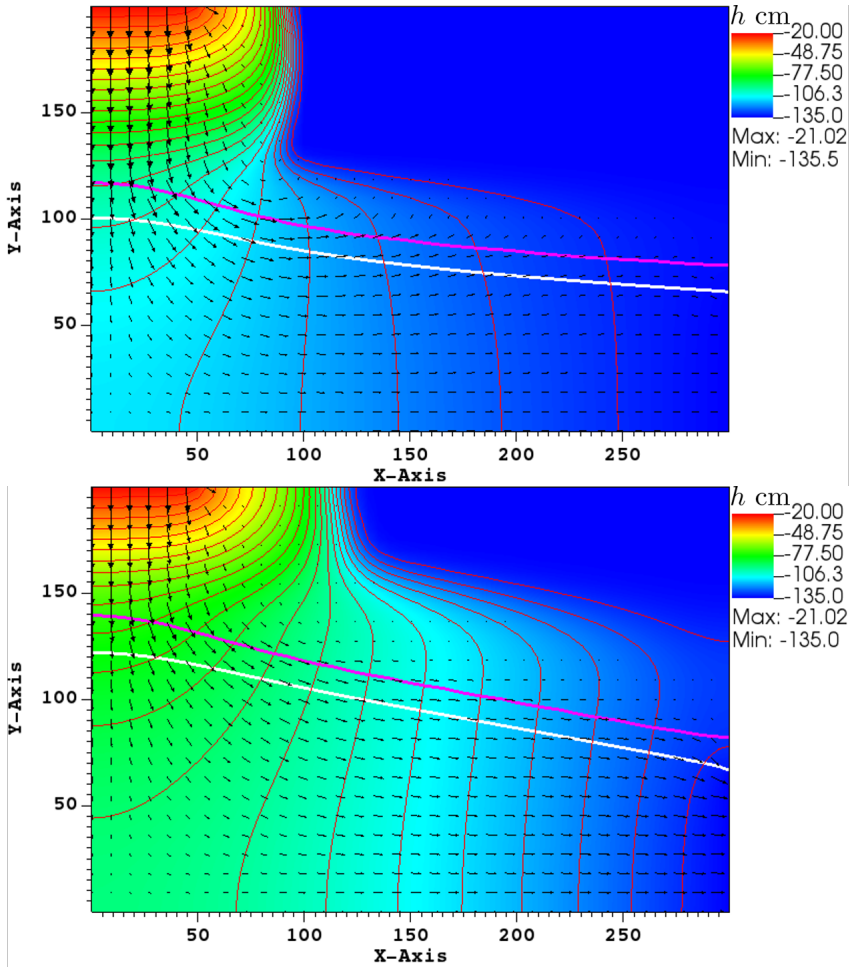


Figure IV.28 – Distribution of hydraulic head h and flux, location of water table (white line) and capillary fringe (magenta line) at $t = 3$ h (top) and $t = 8$ h (bottom). Hydraulic head contours are in red.

Figure IV.28 illustrates a typical field distribution of fluxes and hydraulic head as well as the location of water table and capillary fringe (set at $\theta = 0.29$), at time $t = 3$ h and $t = 8$ h. Hydraulic head values have been turned upside down to correspond with a downward vertical axis and to facilitate comparison with the numerical and experiment results from Vauclin *et al.* [203] provided in Fig. IV.27. At $t = 8$ h, Vauclin *et al.* [203] stated that steady was reached experimentally. Numerically, the present results are not in a steady state because if the simulation is continued, one reaches results from Fig. IV.25. In addition, one should note that the use of different basis functions and approximation order for quadrangles are determinant for the solution quality. Indeed, quadratic monomial basis functions are necessary to get equivalent results to linear Lagrangian basis functions. This is explained by the tensored term.

IV.6.2 Water table drainage

Lowering

Vauclin *et al.* made a drainage experiment of a slab of soil in which experimental details and positions of water table can be found in [201] and [202]. A water table lowering occurs after an instantaneous drop of the imposed hydraulic head. The box is 200 $cm \times 300$ cm. Hydraulic properties are taken

from Vachaud's relations (see Tab. I.1) with $A = 359720.49$, $B = 4.5$, $C = 40000$, $D = 2.9$, $K_s = 40$ cm/h, $\theta_s = 0.3$ and $\theta_r = 0$. This experiment is used as a test-case to show a transient simulation with seepage boundary condition. The domain is initially in hydrostatic equilibrium: $h_0 = 145$ cm. Boundary conditions at the bottom, at the top, on the right and on the left for $z > 145$ cm are $\nabla h \cdot \boldsymbol{n} = 0$ (impervious). On the left, the water level is kept constant in the ditch for $z \leq 75$ cm, $h = 75$ cm. A seepage boundary condition is prescribed for 75 cm$< z \leq 145$ cm. Simulation lasts $T = 20$ h. The configuration is sketched in Fig. IV.29.

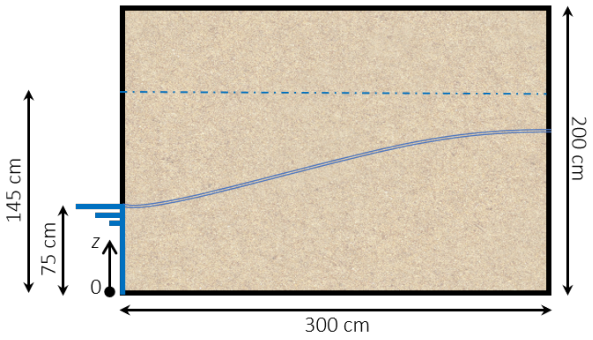


Figure IV.29 – Water table lowering test-case configuration.

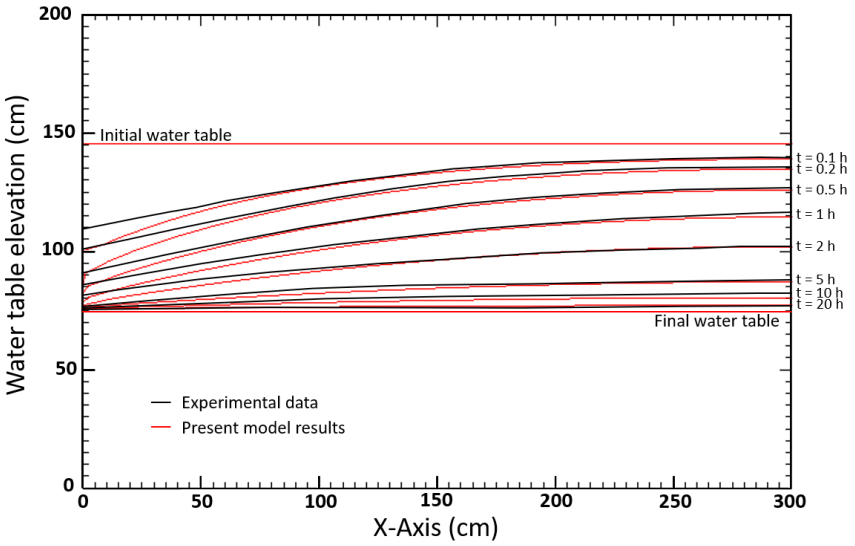


Figure IV.30 – Numerical simulation of two-dimensional lowering with water table position compared to experimental data from Vauclin *et al.* [201].

Numerical simulation is performed with $p = 1$, $\sigma^I = \sigma^D = 100$ on a mesh with $\Delta x = 10$ cm and $\Delta z = 5$ cm. Water table location matches $\psi = 0$. Numerical positions of water table are displayed in Fig. IV.30 together with the experimental data from [201]. Results show that the model is able to reproduce satisfactorily the transient water table dynamics. However, disagreements near the seepage boundary (left) are found between numerical and experimental profiles. Two reasons can explain it. One is in relation with the experimental set-up because water should not pour out completely freely through the experimental device in such a way that seepage boundary condition does not model it properly. Another reason is the determination of hydraulic properties which was done by correlation

from measurements with some fairly important dispersion, due to natural heterogeneity/anisotropy and by neglecting hysteresis [201]. This would explain some discrepancies. In every instance, it is not linked to discretization because results does not change whatever the mesh and order approximation are chosen. Other numerical studies based on this experiment are reported [19, 201, 204]. They also struggle to match water table profiles near the seepage.

Figure IV.31 illustrates a typical field distribution of fluxes and hydraulic head as well as the location of water table and capillary fringe (set at $\theta = 0.29$), at time $t = 0.1$ h and $t = 1$ h.

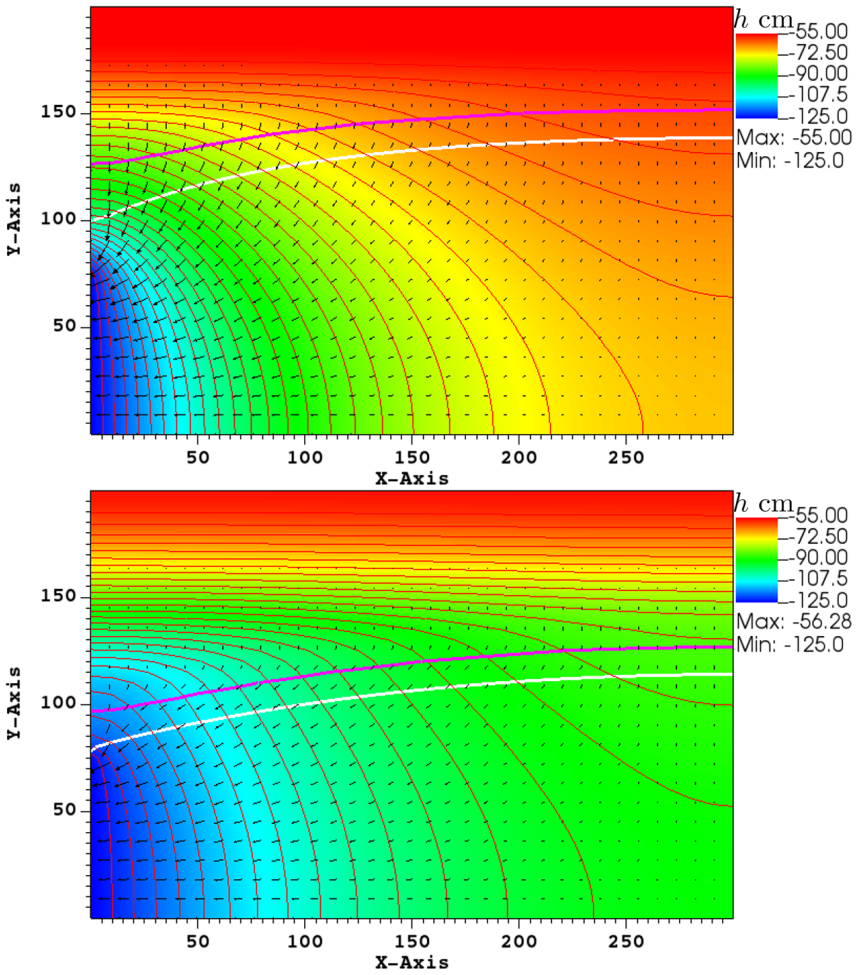


Figure IV.31 – Distribution of hydraulic head h (turned upside down) and flux, location of water table (white line) and capillary fringe (magenta line) at $t = 0.1$ h (top) and $t = 1$ h (bottom). Hydraulic head contours are in red.

Drawdown

Vachaud *et al.* made a drainage experiment of a slab of soil in which experimental details and positions of water table can be found in [200] and [202]. A water table drawdown occurs after an instantaneous drop of the imposed hydraulic head. The box is 200 cm × 300 cm. Hydraulic properties are taken from Vachaud's relations (see Tab. I.1) with $A = 359720.49$, $B = 4.5$, $C = 40000$, $D = 2.9$, $K_s = 40$ cm/h, $\theta_s = 0.3$ and $\theta_r = 0$. This experiment is used as a test-case to show a transient simulation with seepage boundary condition. The domain is initially in hydrostatic equilibrium: $h_0 = 143$ cm. Boundary conditions at the bottom, at the top, on the right for $z > 143$ cm and on the left for

$z > 143$ cm are $\nabla h \cdot \boldsymbol{n} = 0$ (impervious). The water level is kept constant in the right ditch for $z \leq 143$ cm, $h = 143$ cm, and, in the left ditch for $z \leq 80$ cm, $h = 80$ cm. A seepage boundary condition is prescribed for 80 cm$< z \leq 143$ cm. Simulation lasts $T = 10$ h. The configuration is sketched in Fig. IV.32.

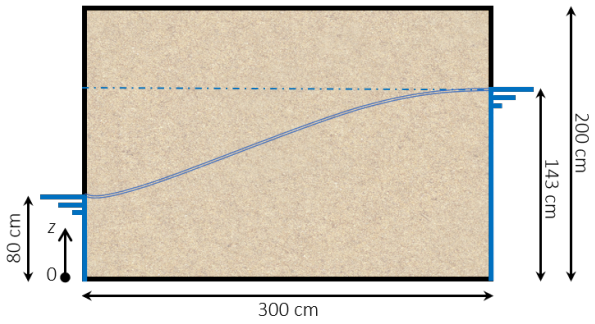


Figure IV.32 – Water table drawdown test-case configuration.

Numerical simulation is performed with $p = 1, \sigma^{\mathrm{I}} = \sigma^{\mathrm{D}} = 100$ on a mesh with $\Delta x = \Delta z = 10$ cm. Water table location matches $\psi = 0$. Numerical positions of water table are displayed in Fig. IV.33 together with the experimental data from [200]. Results show that the model is able to reproduce satisfactorily the transient water table dynamics. However, discrepancies are observed between numerical and experimental profiles, in particular near the seepage boundary (left) with a small late for numerical results. The same explanation than the previous can be made.

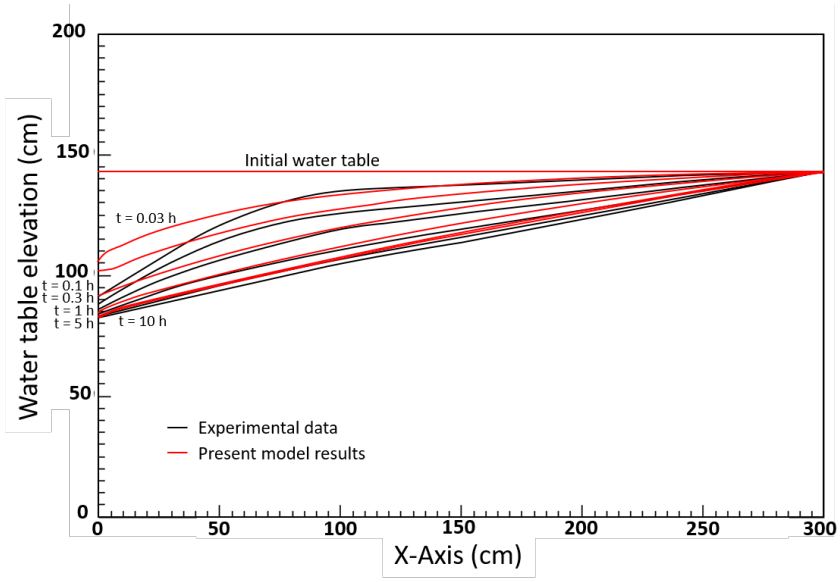


Figure IV.33 – Numerical simulation of two-dimensional drawdown with water table position compared to experimental data from Vauclin *et al.* [200].

Figure IV.34 illustrates a typical field distribution of fluxes and hydraulic head as well as the location of water table and capillary fringe (set at $\theta = 0.29$), at time $t = 1$ h.

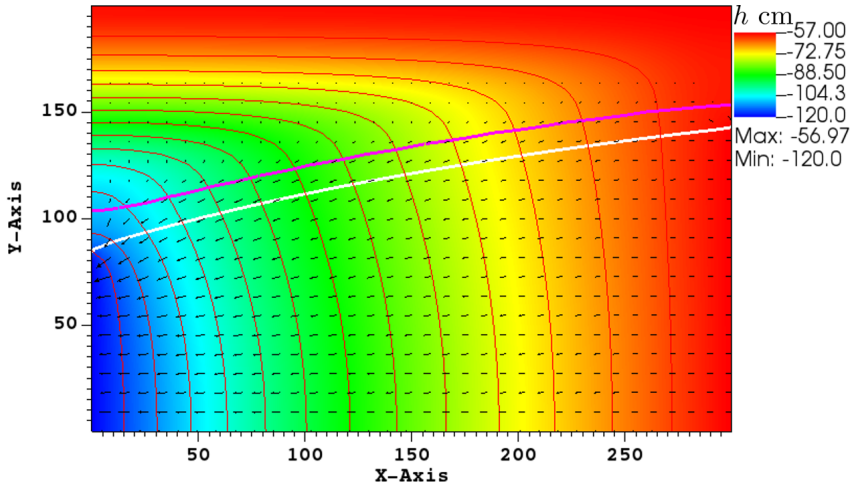


Figure IV.34 – Distribution of hydraulic head h (turned upside down) and flux, location of water table (white line) and capillary fringe (magenta line) at $t = 1$ h. Hydraulic head contours are in red.

IV.7 Assessment of La Verne dam wetting

IV.7.1 Experimental configuration

The model is here used to simulate the full-scale wetting of a multi-materials dam: the La Verne dam. La Verne dam was constructed in 1991 to supply Gulf of Saint-Tropez (south of France) with drinkable water. The dam is 42 m high and peaks at 90 m above sea level. Its width ranges from -102.3 m to 110.5 m. The reservoir is on the left in Fig. IV.35. The genuine La Verne dam inner section shows a complicated heterogeneous material patterns, partially described in [81, 209]. The dam is an earth-filled embankment dam. A clay core allows for impermeability while outer zones are semi-permeable alluvial shells. The dam is protected by rip-rap upstream. Filters and drains of very permeable gravel-like materials secure the dam inner saturation. The dam foundation lies on an impervious rock stratum thanks to concrete injections. A fine loam layer with vegetation covers the downstream dam part. The filling of La Verne dam was controlled with instruments [209] during 40 days, which provides a unique and challenging dataset to test the model. In particular, the reservoir height evolution is available as well as three groundwater hydraulic head measurements from the core, see Fig. IV.35 and IV.38. The three sensors are at $z = 55$ m and their x-coordinates are -10.5 m, -3 m and 5 m.

For the simulation, the reservoir height is used as a forced boundary condition in the form of a Dirichlet boundary condition monitored by a function based on experimental data (see Fig. IV.36). Since comparison focusses on the core of the dam, it is assumed that outer high permeable and/or fine layers materials do not have a significant impact on inner groundwater flows and saturation. The downstream thin loam layer, the filter-isolated toe drain and the lower submerged rip-rap are therefore not represented by the simulation and assimilated to the adjacent materials (see Tab. IV.3 for description of numerical dam structure). A Dirichlet boundary condition for hydraulic head is prescribed at the downstream boundary to model remaining water in the filter which is diverted to exit horizontally. The upper rip-rap and the cofferdam are simulated as such to assess the method robustness owing to their direct contact with the upstream dynamic boundary condition (forcing) and their hydraulic properties are steep. Rock/concrete foundations are supposed to be perfectly impervious which means a zero-valued Neumann boundary condition can be prescribed. Finally, seepage boundary condition are prescribed everywhere else because outflow can drain from the exposed upstream and downstream

shell slopes. The actual simulated configuration of La Verne dam is sketched in Fig. IV.35. Five different materials are simulated (Tab. IV.3), with hydraulic properties provided by Bonelli *et al.* [81]. Their quotation marks designation in Tab. IV.3 does not reflect the actual soil texture but can serve as a guide. The dam wetting simulation involves the main tools presented in this thesis for the adaptive strategy: the WDG method in combination with the AMR technique driven by the estimation-based error indicator. Mesh adaptation is done every five time steps. Fixed-point iteration is used together with adaptive time stepping. Order of approximation is quadratic for space and one for time in order to have a robust and cost-effective computation. Penalty parameters are $\sigma_{\mathrm{I}} = \sigma_{\mathrm{B}} = 100$.

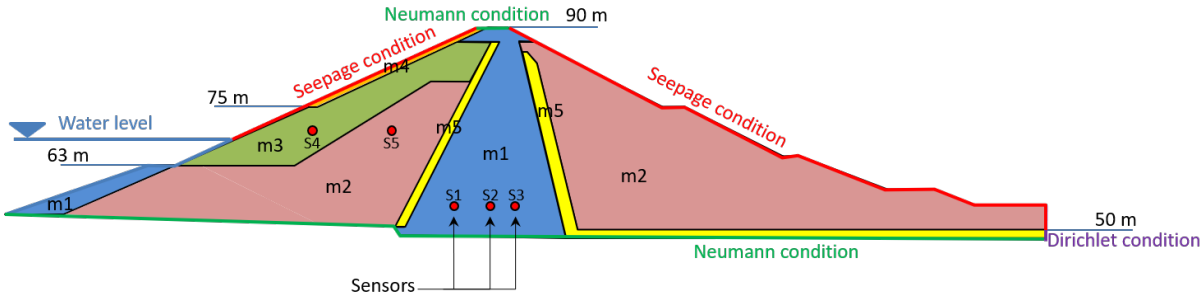


Figure IV.35 – Geometry, materials and boundary conditions of La Verne dam for the numerical case. S1, S2, S3 are numerical sensors corresponding to the experimental sensors while S4 and S5 are additional purely numerical sensors.

Materials	Components	Constitutive laws	θ_{s}	θ_{r}	$\mathbb{K}_{\mathrm{s}}(1,1)\left(m \cdot d^{-1}\right)$	$\mathbb{K}_{\mathrm{s}}(2,2)\left(m \cdot d^{-1}\right)$	Specific parameters
m1 "clay"	Core Cofferdam	Van Genuchten-Mualem	0.23	0	2.592×10^{-3}	8.64×10^{-4}	$\alpha = 0.08m^{-1}, n = 1.2, l = 0.5$
m2 "silty alluvium"	Upstream shell Downstream shell	Van Genuchten-Mualem	0.25	0	5.184	1.728	$\alpha = 0.01m^{-1}, n = 2.1, l = 0.5$
m3 "sandy alluvium"	Outer upstream shell	Vachaud	0.22	0	5.184	1.728	$A = 2.99 \times 10^{-4}m^{5}, B = 5$ $C = 6.34 \times 10^{-2}m^{2.9}, D = 2.9$
m4 "rocks"	Protection rip-rap	Vachaud	0.27	0	17.28	8.64	$A = 2.99 \times 10^{-4}m^{5}, B = 5$ $C = 6.34 \times 10^{-2}m^{2.9}, D = 2.9$
m5 "gravel"	Filters Drains	Vachaud	0.32	0	17.28	8.64	$A = 2.99 \times 10^{-4}m^{5}, B = 5$ $C = 6.34 \times 10^{-2}m^{2.9}, D = 2.9$

Table IV.3 – Materials of La Verne dam

The simulation of La Verne dam case is a very challenging benchmark for numerical models, combining heterogeneous materials, steep constitutive laws and dynamic boundary conditions. However, despite of its interest, the La Verne dam benchmark remains only partly documented by in-situ instruments, precluding a comprehensive quantitative confrontation between observations and model. In particular, the hydraulic properties, which have a drastic effect on flow dynamics, remain approximative due to the absence of direct in-situ characterization. In addition, note that no initial in-situ values for water table elevation and saturation are available inside the dam, which can also strongly affect the subsequent evolution. For the present simulation, the initial water table is imposed at 67 m to fit the initial values of experimental hydraulic head. Finally, Bonelli *et al.* pointed out that hydromechanical coupling should be considered to be fully consistent with the real case, see [209].

IV.7.2 Numerical results

Figure IV.36 depicts hydraulic head (color contours), water table (white line) and adaptive mesh for both simulations every 10 days. Number of elements, magnitude of time steps and the value of global

error estimate during the simulation are shown in Fig. IV.37. Time evolution of simulated hydraulic head and water content at the three experimental sensors (S1 to S3) and the two additional numerical sensors (S4 and S5) are displayed in Fig. IV.38.

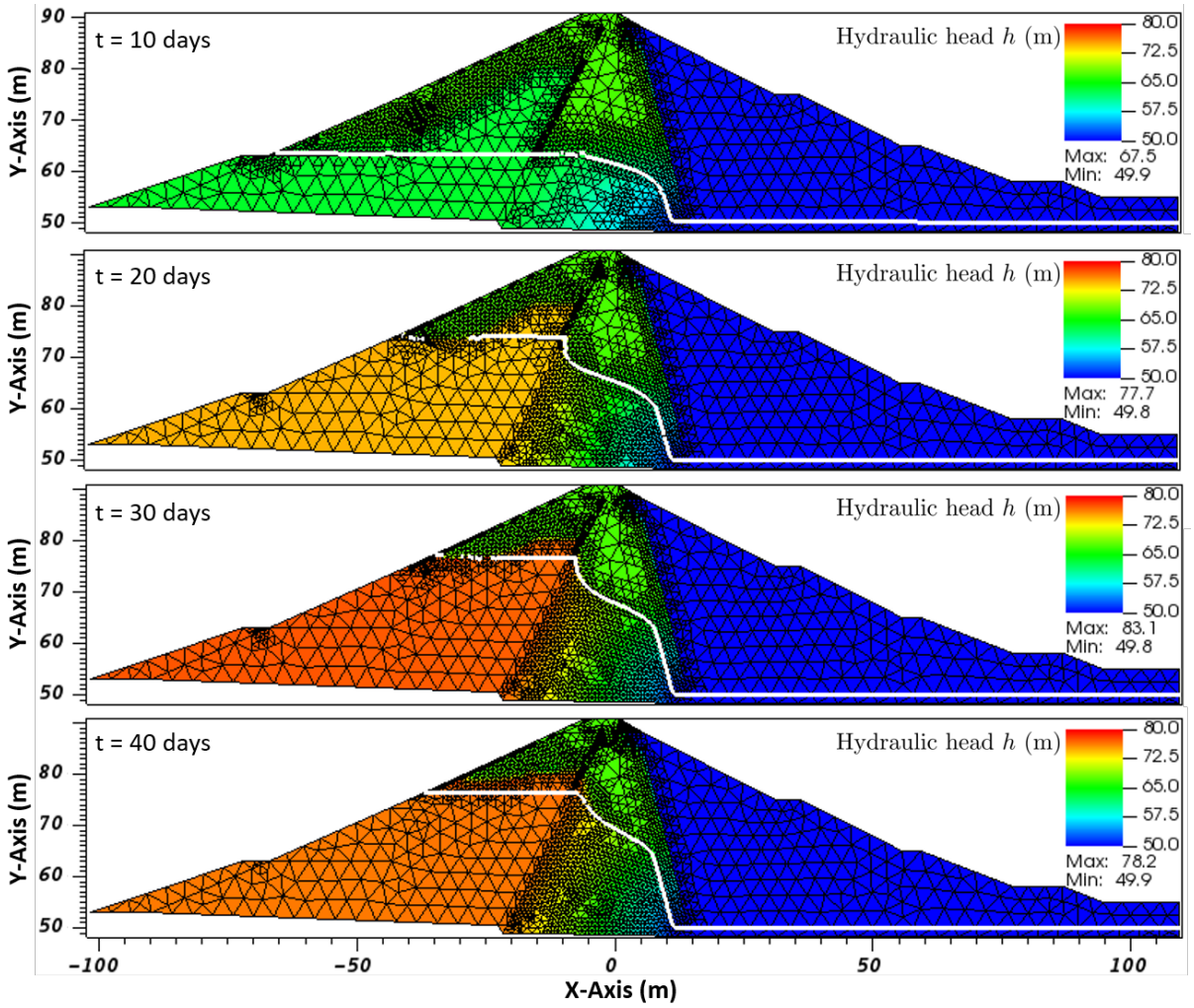
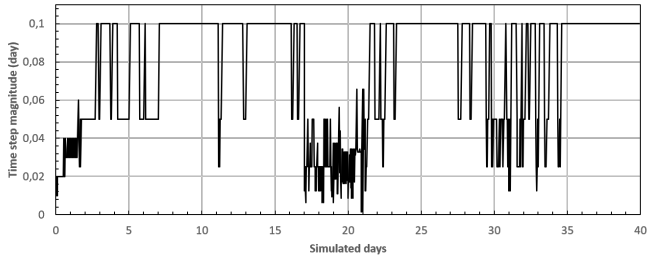


Figure IV.36 – Time evolution of hydraulic head h, mesh and water table (white line) during La Verne dam wetting.

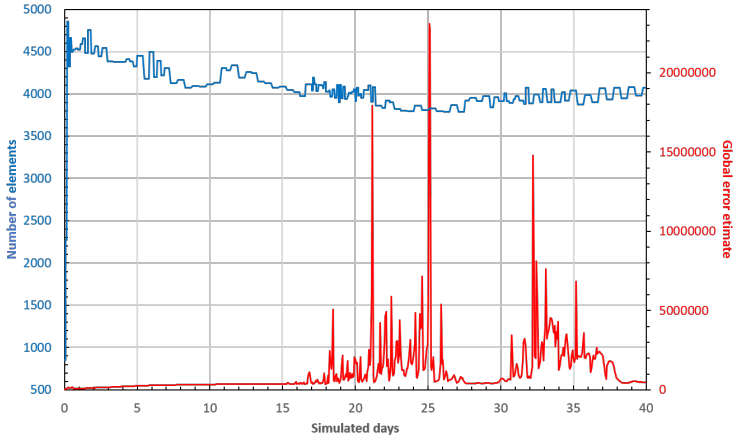
It is recalled that initial reservoir free surface and water table elevation are at 57 and 67 m, respectively. Note that the water table is not displayed in Fig. IV.36 at some points due to insufficient resolution and oscillations w.r.t the front sharpness, which are detailed in Section IV.7.3 thereafter. The numerical model is observed to provide a good overall description of the dam wetting process. Figure IV.36 shows that the propagation of the wetting wave inside the dam is strongly dependent on the inner materials properties.

Two major stiff periods for the simulation can be noticed in Fig. IV.37: one is approximately between 17 and 21 days and the other between 29 and 35 days. They match a decrease of the time stepsize showing rapid variations. This means nonlinear iterative process struggles to converge. After growing a lot at the very beginning, the number of elements stabilizes around 4100 elements during the simulation despite the global error estimate is increasing for the two stiff periods. This is because mesh refinement is unable to control errors where it is needed because the maximum refinement level has already been

reached in these regions. Usually, it comes with the development of overshoots/undershoots which are discussed later.



(a) Time step evolution



(b) Number of elements and global estimate evolution

Figure IV.37 – Evolution of some numerical variables during La Verne simulation.

The first eleven days are characterized by an adjustment phase during which the reservoir surface remains below the initial water table elevation. In Fig. IV.37, a decrease in time stepsize is observed when the reservoir surface hits the initial water table elevation and starts to saturate dry regions. Hydraulic head and water content at numerical sensors S4 and S5 show a very contrasting behaviour (Fig. IV.38b and c). This reveals that, even if both m2- and m3-materials show the same hydraulic conductivity at saturation, the difference in constitutive laws induces a radically different dynamics of the capillary fringe. At S4, the head remains nearly constant during the adjustment phase while the water content is very low. The m3-material is here nearly desaturated, well above a thin capillary fringe. By contrast, S5 is within the large capillary fringe associated to the m2-material, providing a much higher water content. The hydraulic head shows first a small decay, probably due to the drain proximity, before starting to rise in response to the reservoir filling. Around day 14, both S4 and S5 sensors show a simultaneous regular rising, slightly lagged from the reservoir level. A strong jump is observed at S4 (Fig. IV.38c), indicating the crossing of a sharp capillary fringe. While rising, the water table remains nearly horizontal in the upstream m2- and m3-materials zone but a sharp front is observed to develop across the left inner drain (m5-material), see Fig. IV.36. Small numerical oscillations in hydraulic head are observed in Fig. IV.38 but do not affect the global dynamics. Further downstream and lower within the dam, the hydraulic head simulated at S1, S2 and S3 shows the slow wetting wave propagation, with attenuation and damping (Fig. IV.38b). At S1, S2, S3, the water content remains constant during the whole simulation, in accordance to the sensors position below water table.

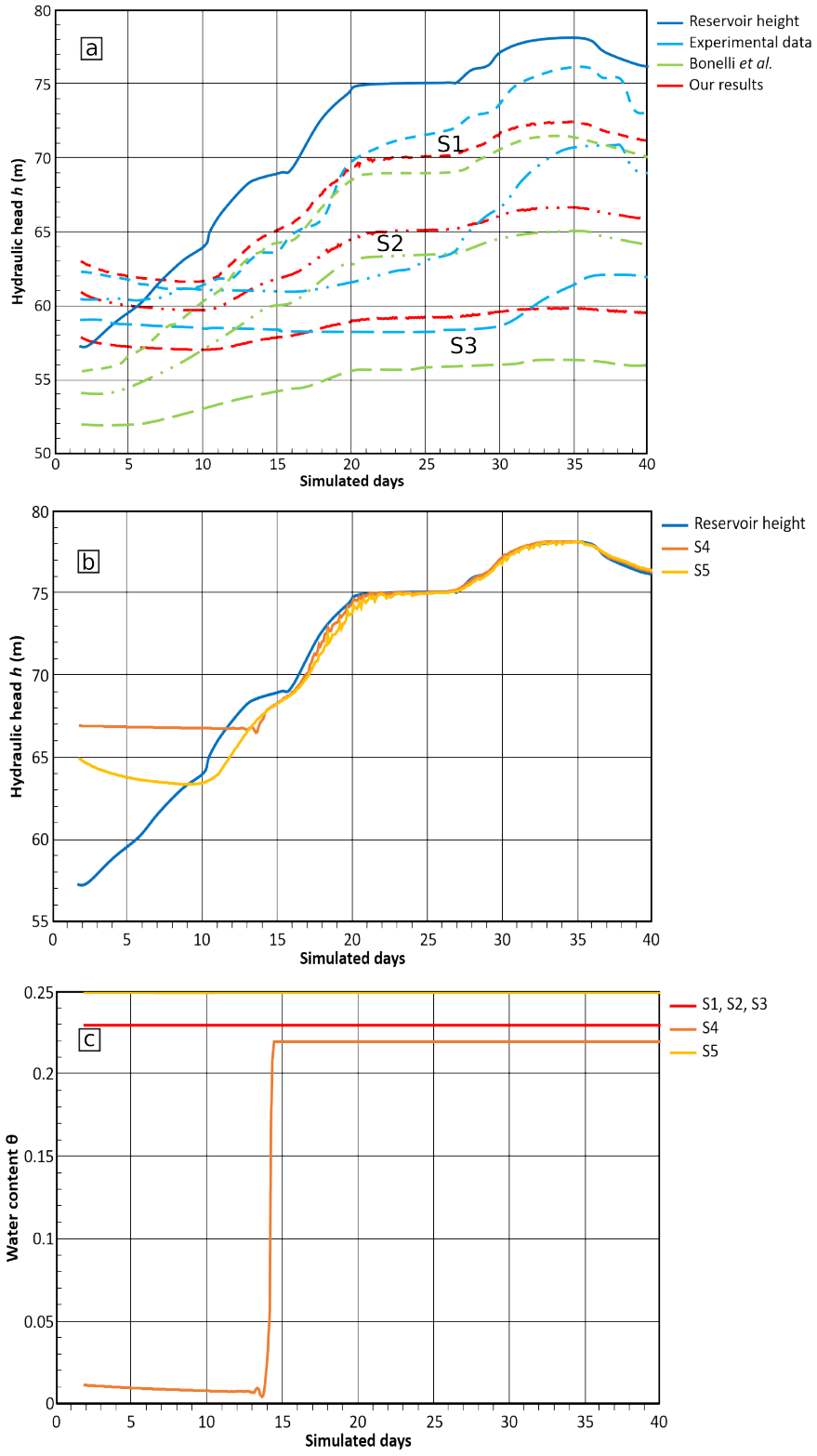


Figure IV.38 – Time evolution of inner dam hydraulic parameters. a) Hydraulic head at sensors S1 (short-dashed lines), S2 (dash-dotted lines) and S3 (long-dashed lines) from the present simulation (red), the experimental data (light blue) and the previous results of Bonelli *et al.* [81]. The reservoir height is depicted in dark solid line. b) Hydraulic head at numerical sensors S4 (orange) and S5 (yellow). c) Water content at S1 to S5 sensors (note that S1, S2 and S3 show similar values).

When compared to hydraulic head experimental data (see Fig. IV.38a), the overall time evolution and order of magnitude are well represented. Some discrepancies are observed, in particular the late increase of head which is experimentally observed on each sensor around 30 days, which tends to be underestimated by the model. However, recalling the lack of experimental control evoked before, more detailed comparison should be made with caution. For the sake of comparison, the previous numerical results presented by Bonelli *et al.* [81] are recalled in Fig. IV.38 in green lines. While the overall time evolution is rather similar between models, a better agreement with experimental data is obtained with the present simulation.

IV.7.3 Discussion and investigation

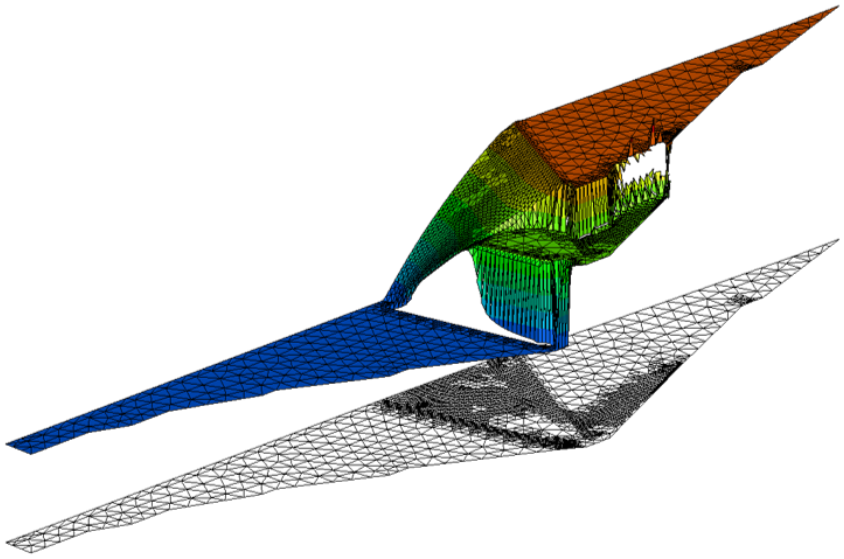
The proposed modelling strategy for Richard's equation showed very satisfactory performance in reproducing analytical and laboratory test cases and in simulating complex full-scale experiment. In this latter case, the model is able to capture automatically wetting fronts which are moving dynamically thanks to AMR. In particular, the fronts associated with heterogeneity are well resolved by the WDG technique. This is illustrated by the discontinuous water content distribution in Fig. IV.40. Moreover, the simulation is robust since it is able to compute completely without user intervention and handle many features: degeneracy for saturated/unsaturated zones, dynamic forcing boundary condition and seepage. However, the simulation struggles to capture wetting fronts associated with nonlinear varying diffusivity resulting in severe overshoots/undershoots for the transition between saturated and unsaturated zones in very permeable materials. To have a better idea of the situation, Fig. IV.39 shows an elevated-view of hydraulic head at $t = 30$ days. The view illustrates the jumps in the solution introduced by the WDG method, the adapted mesh and the aforementioned overshoots.

Further insight on this behaviour can be discussed using the Péclet number framework. Richards' equation is an elliptic-parabolic equation predominantly diffusive, it shares properties with convection-diffusion equation and can behave like them. For example, sharp wetting fronts and internal layers can be reinterpreted within this context. The Péclet number represents the ratio of the rate of advection to the rate of diffusion. This dimensionless number is useful to analyse locally the nature of the flow regime and determine whether advection or diffusion is dominant. Following ideas for subsurface flow from Huyakorn and Pinder [210] or remarks from [90], the oscillations in the solution for Galerkin finite element method are linked to the Péclet (and Courant) number. Sharp fronts carry short wavelength harmonics of the Fourier series not recognised by the numerical scheme for the current discretization [211]. El-Kadi and Ling [195] derive a Péclet number and a Courant number for Richards' equation in 1D to estimate the size of spatial and temporal local steps for acceptable accuracy of the numerical simulation. However, the Péclet number definition is slightly different from Eq. (IV.22). A recent study tries to select suitable approximation for Richards' equation based on the Péclet number [212].

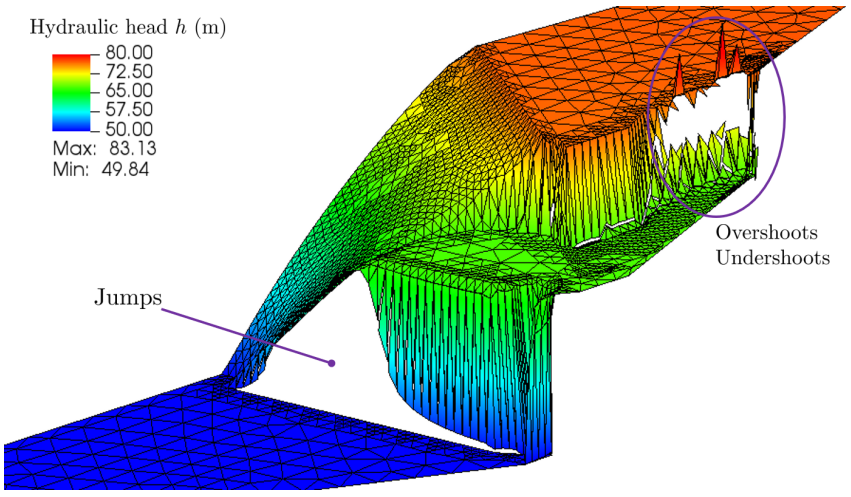
For this study, the Péclet number is extracted from the saturation-based formulation of Richards' equation given in Eq. (I.49). The latter is similar to a convection-diffusion equation with the water content θ as single unknown. This is a Fokker-Planck equation in essence [213] whose form is useful to study stability and investigate travelling front solutions of wetting and drying fronts [83]. This formulation is undefined in saturated zones because hydraulic diffusivity $\mathbb{D}(\psi) = \dfrac{\mathbb{K}(\psi)}{C(\psi)}$ tending to infinity but this zone is not of interest for the Péclet number since Richards' equation is purely diffusive. Moreover, it is assumed that $\theta \neq 0$. This assumption is not restrictive because if water content drops to zero, nothing happens and, in practice, there is often a residual water content. The total flux can be separated into a diffusive flux defined by hydraulic diffusivity and driven by capillarity,

and an advective flux defined by hydraulic conductivity/water content and driven by gravity. The saturation-based formulation of Richards' equation given in Eq. (I.49) can be rewritten as:

$$\partial_t \theta - \nabla \cdot \left(\mathbb{D}(\theta) \nabla \theta + \frac{\mathbb{K}(\theta) \nabla z}{\theta} \theta \right) = 0. \tag{IV.21}$$



(a) Whole domain view



(b) Close-up view

Figure IV.39 – Elevated-view of hydraulic head h with mesh for La Verne dam wetting at $t = 30$ days. Point of view is from above, downstream.

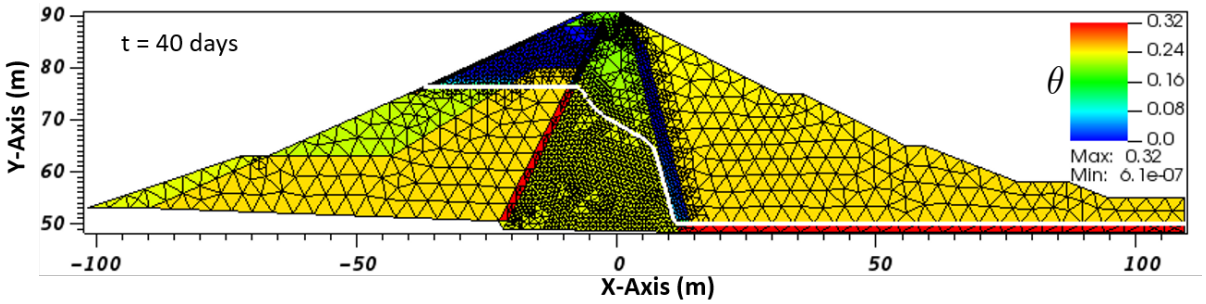
The Péclet number can therefore be defined as:

$$\mathrm{Pe} = \frac{\|\mathbb{K}(\theta)\| L}{\|\mathbb{D}(\theta)\| \theta} = \frac{C(\theta) L}{\theta}, \tag{IV.22}$$

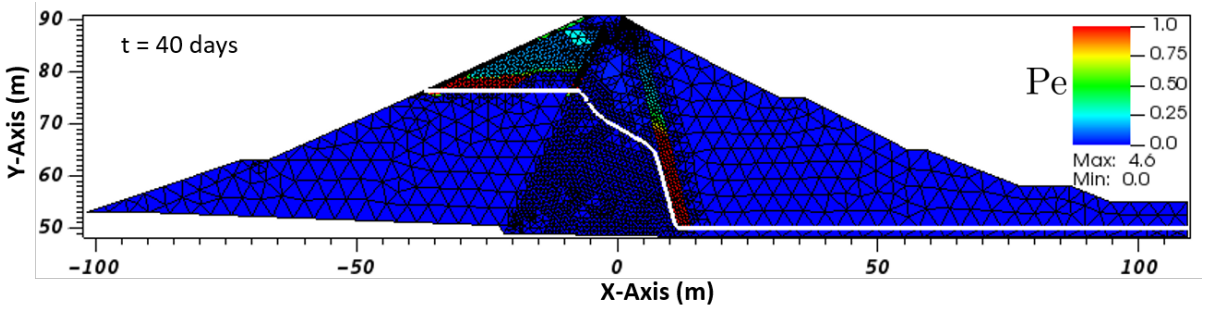
with L the characteristic linear dimension assimilated to an element size h_E.

High Péclet numbers, from 1 to more than 5, are clearly reached in the sharp-gradient areas in Fig. IV.40, with m3-, m4- and m5-materials (*i.e.* drains, rip-rap and upstream shell). Consequently, water table

is not straight anymore and a delay in elevation is observed upstream because of over/undershoots, see Fig. IV.36. In this case, Péclet number can be of interest to identify very high permeable regions induced by hydraulic properties where advection dominates diffusion and mesh resolution strains to capture the flow pattern.



(a) Distribution of water content at $t = 40$ days

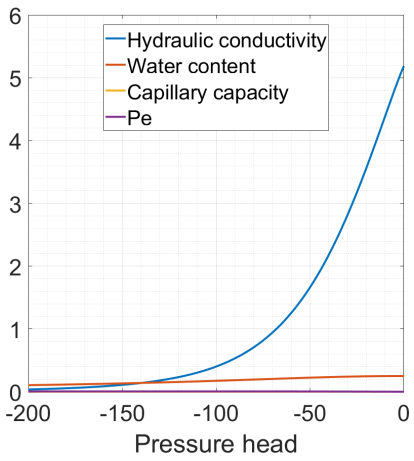


(b) Distribution of Péclet number at $t = 40$ days

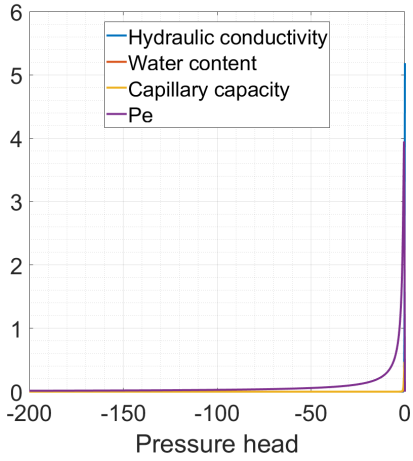
Figure IV.40 – Heterogeneity and nonlinear varying diffusivity highlighted respectively by water content and Péclet number for La Verne dam simulation with the reference simulation.

Figure IV.41 illustrates several hydraulic properties for m2- and m3-materials in the horizontal direction. m3-Material holds stronger gradients than m2-material in the vicinity of zero (capillary fringe). In particular, hydraulic diffusivity $\mathbb{D}$ shows a singularity for $\psi \longrightarrow 0^-$. Besides, the Péclet number is higher for m3-material with a maximum corresponding to the foot of hydraulic conductivity.

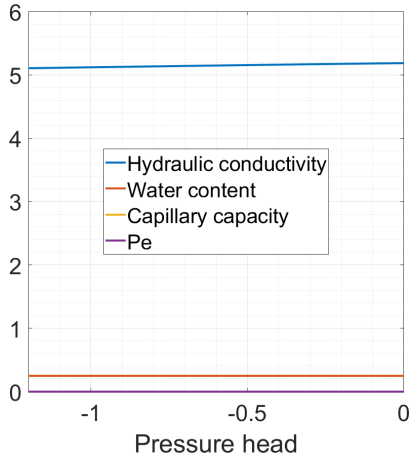
In order to overcome these limitations, an additional ad-hoc configuration, named *augmented* simulation, has been implemented. This one does not use WDG framework and other difference lies in the mesh and tolerances for nonlinear convergence criterion. Finer discretization is allowed in permeable material that is to say the m1-, m2- and m3-materials. The mesh is also refined around water table and both for the gradient-based and estimation-based error indicators. Instead of the reference simulation, 4-order BDF is employed and hydraulic properties are relaxed with a numerical minimal value to avoid complete degeneracy. The *augmented* computation allowed to eliminate spurious oscillations from the solution thanks to a more refined mesh, which explains why WDG method is not needed. As the mesh is extremely refined, the *augmented* simulation takes approximately 13.5 times longer to compute (42 h 37 min) than the reference simulation (3 h 11 min). One core Intel(R) Xeon(R) CPU 5-2630 v3 with 2.40 GHz was used. Figure IV.42 depicts hydraulic head (color contours), water table (white line) and adaptive mesh for both simulations every 10 days for the *augmented* simulation to compare with the reference simulation in Fig. IV.36.



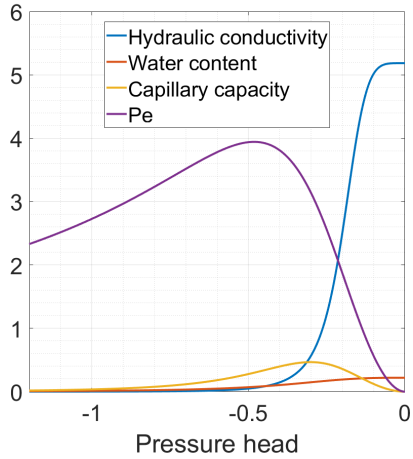
(a) Hydraulic properties for m2-material



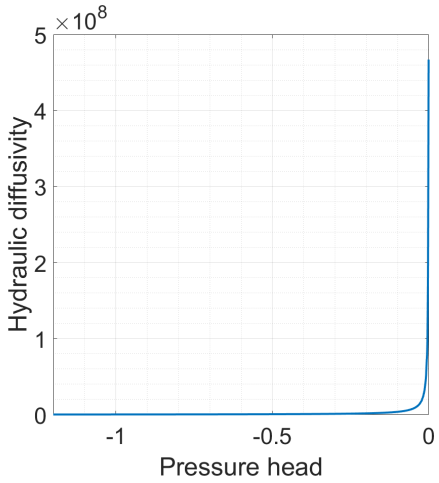
(b) Hydraulic properties for m3-material



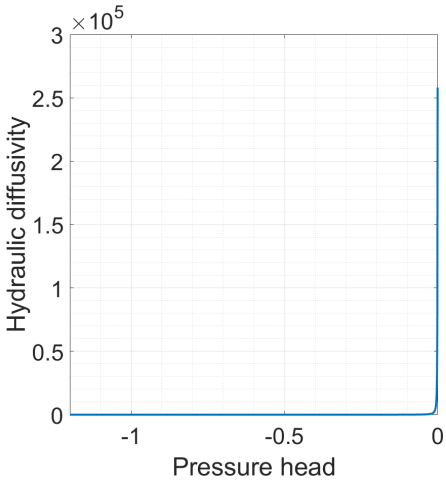
(c) Hydraulic properties for m2-material (zoom)



(d) Hydraulic properties for m3-material



(e) Hydraulic diffusivity for m2-material (zoom)



(f) Hydraulic diffusivity for m3-material

Figure IV.41 – Hydraulic properties functions and Péclet number (with $L = 1$) for m2- and m3-materials (Van Genuchtem-Mualem and Vachaud's relations respectively) for La Verne dam in the horizontal direction.

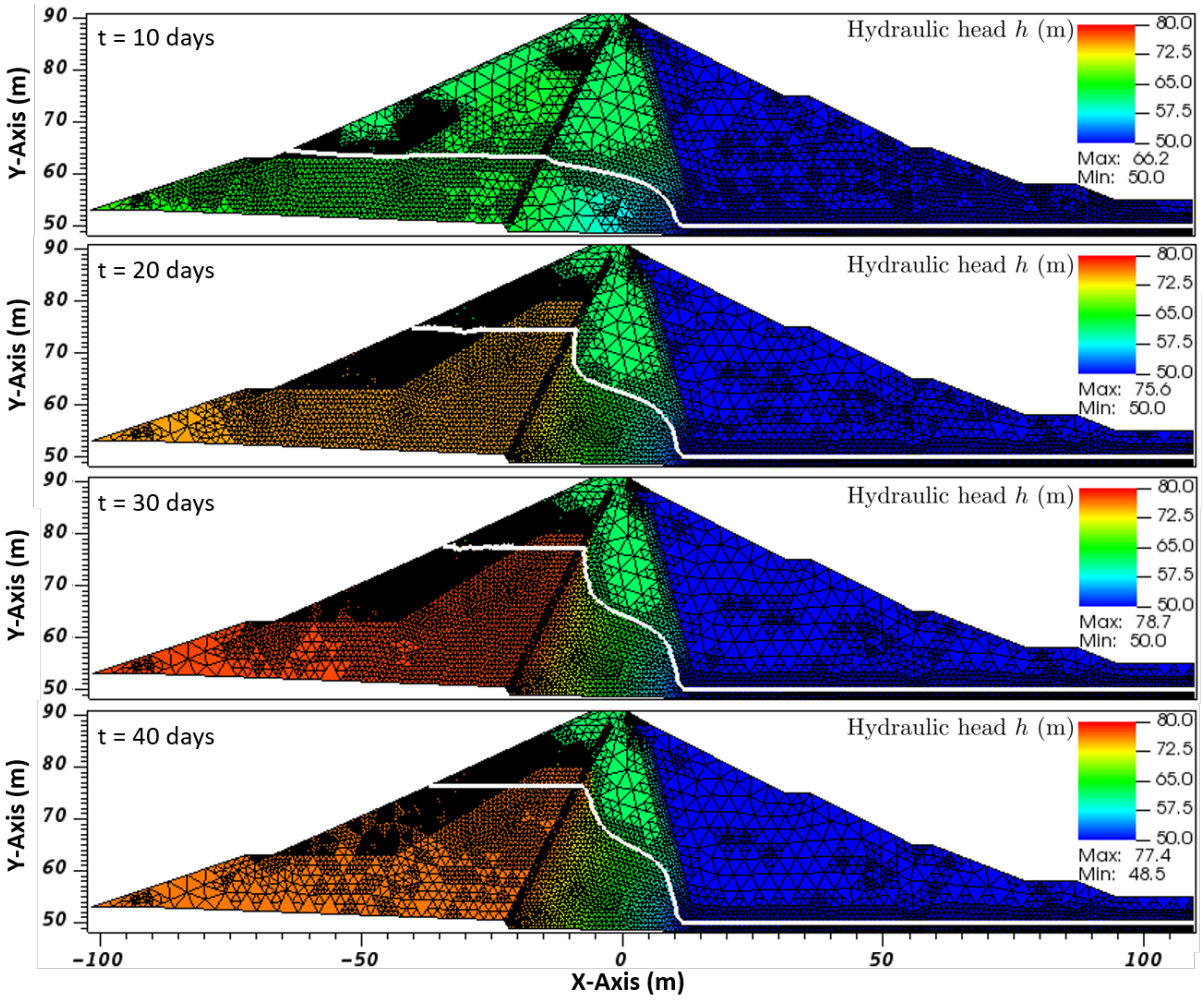
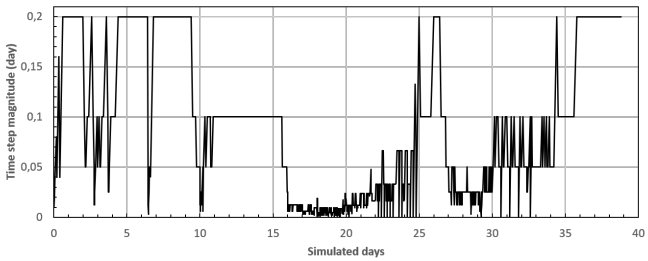


Figure IV.42 – Time evolution of hydraulic head h, mesh and water table (white line) during La Verne dam wetting for the *augmented* simulation.

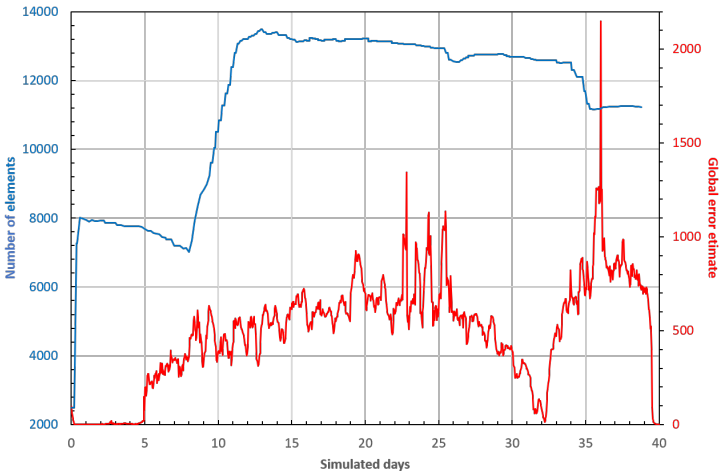
Figure IV.43 illustrates the issues of La Verne dam simulation with the *augmented* computation. Time step is decreasing at 10 days when water table meets the m3-material. Two drastic reductions happen next: around 15-20 days and 28 days. This corresponds to the water table moving through m4-material and reaching the filters made of m5-material. The solver struggles in these conditions to reach convergence which makes time step decrease drastically. At the end of the simulation, time step is increasing largely which matches drainage with the drop of the reservoir level. One can also remark that the number of elements is increasing at the beginning of the simulation in the course of successive refinements. The errors are controlled which is shown by the global estimator remaining quite low. However, as soon as the specified maximum level of refinement is reached within a block, the error is increasing freely. It would be interesting to attempt different process for the selection of elements to be refined/coarsened.

The simulation of La Verne dam filling shows many difficult points which lead to two unwanted effects. The first one is the development of non-physical oscillations around steep wetting fronts associated with nonlinear varying diffusivity which affect solution accuracy. The other one is nonlinear convergence possible only with very small time steps so that the calculation duration is high. These challenges illustrate numerical issues of Richards' equation for heterogeneous media with high-varying diffusivity. The improvements offered by the *augmented* simulation support the idea that a series

of heuristic tools (threshold values, maximum refinement level, refinement frequency, adaptive time steps) should be investigated to improve accuracy while maintaining cost-effectiveness.



(a) Time step evolution



(b) Number of elements and global estimate evolution

Figure IV.43 – Evolution of some numerical variables during La Verne simulation with *augmented* computation.

Summary

The work done in this chapter has showed validations for the *Rivage* code as well as numerical results supporting the proposed solving strategy. A series of 1D infiltration in a soil column illustrates behaviour of Richards' equation solution according to different relations for the hydraulic properties. It has been seen that wetting fronts induced by the hydraulic properties may be sharp and move fast causing spurious oscillations which are not physical. Global mass balance is respected but element-wise mass balance should be assessed more when oscillations occur.

To deal with these instabilities, the proposed strategy is promising: AMR is capturing the moving fronts while the WDG framework accounts for jumps in the solution to better approximate the fronts. Thanks to Tracy's benchmark, the error indicator based on error estimation is chosen for its ability to drive mesh adaptation in synergy with the weighted approximation. It produces cost-effective mesh compared to the gradient-based error indicator. Besides, the subpart od the estimation-based error indicator suggest a way to make *hp*-adaptation in future studies.

Finally, *Rivage* code is examined for 2D unsaturated porous media flow with laboratory and large-scale experiments. First of all, a set of water table recharge and drainage validate the model for saturated-unsaturated conditions and complex flow dynamics involving the seepage boundary condition. In particular, the adaptive time stepping is mandatory to get convergence robustness. Last, experimental data from La Verne dam wetting after reservoir impoundment provide a challenging case to simulate. Indeed, the earth-filled dam has a multi-material structure and the reservoir filling induces complex dynamics with multiple wetting fronts, dynamic forcing and seepage. The proposed strategy is very satisfactory to reproduce experimental data and computational time stays reasonable. Moreover, sharp internal layers due to material heterogeneity are well resolved. However, the simulation struggles to capture moving wetting fronts associated with nonlinear varying diffusivity resulting in severe overshoots/undershoots. The Péclet number gives interesting information: it happens at transition between saturated and unsaturated zones in very permeable materials. A more refined simulation of La Verne dam wetting suggests that the strategy can be improved by investigating heuristic tools on which adaptation relies.

Chapter V

Applications to swash groundwater dynamics in sandy beaches

"La mer, la mer, toujours recommencée !"
— Paul Valéry, *Le cimetière marin*, 1920

Groundwater dynamics in sandy beaches is a matter of interest since beach represents interface between land and open sea and accounts for a wide range of processes: biogeochemical cycles, sediment transport, diffusion of dissolved materials, fresh/salt water exchanges between ocean and coastal aquifer, *etc*. It is therefore directly involved in the anthropogenic issues in nearshore areas in the global context of climate change and growing urbanization: beach erosion, submersion, aquifer salinization, *etc*. Many laboratory and field studies have been devoted to understand groundwater flows inside sedimentary beaches [3, 5]. Observational approaches are however still limited to resolve the full spatio-temporal variability of governing variables, in particular the fields of head, saturation and flux. Numerical modelling stands out as a promising prospect, allowing to evaluate more accurately water table fluctuations, flow velocities or infiltrations/exfiltrations during swash events which are for now very difficult to measure directly. Nevertheless, up to now, few models have been devoted to the simulation of wave-driven groundwater flows and their scope remains limited [9, 214].

The main purpose of this chapter is to present the application of *Rivage* code and models described in the previous chapters for beach groundwater in the swash zone. The action of waves is modelled by a weak coupling with the dynamic forcing boundary condition described in Section I.5. First of all, a synthetic description of the swash zone is given highlighting the roles of waves on groundwater. An application aims to see if the code is able to simulate the main groundwater features on an idealized beach case. Additional cases are under consideration to look further into swash groundwater dynamics. They are under construction and analysis.

V.1 Wave-forced beach groundwater dynamics

V.1.1 Waves hydrodynamic forcing

Several large families of surface waves can be considered in the coastal zone. Roughly speaking, it can be distinguished:

- the gravity waves of period included in 3 and 20-22 s. There are two sub-categories: the swell for periods higher than 6 s, the sea of the wind below;
- the tide and its harmonics, with typically diurnal or semi-diurnal periods;
- in between, it is the domain of long waves, which include in particular seiche and infragravity waves (between 20 and 500 s);
- surge waves linked to variations in atmospheric pressure, with day to week periods.

The swash zone is a dynamic region subject to wave action, see Fig. V.1. A swash event shows generally two basic phases: uprush (onshore flow) and backwash (offshore flow). An example is given by Fig. V.2. Generally, uprush has higher velocity and shorter duration than backwash. Uprush starts with great velocities diminishing until its end, whereas backwash velocity increase towards its end.

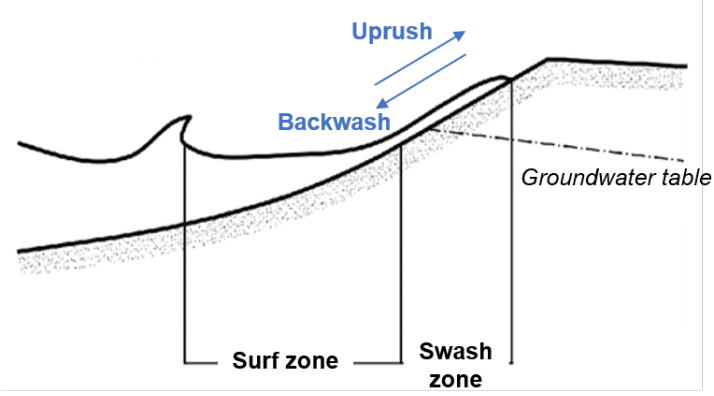


Figure V.1 – Sketch of the swash event consisting in uprush and backwash.

(a) Uprush (b) Backwash

Figure V.2 – Swash on Slea Head beach, Ireland.

The dynamics of the interaction between surface waves and the beach depends on the forcing conditions (type of waves: period, amplitude) the slope of the beach and the grain size of the sediment. In nature, the beach morphology will adapt to the hydrodynamic forcing conditions and the grain size, whose spatio-temporal distribution can also adapt to the forcing. For the present applications, it will first be assumed that grain size is homogeneous (constant and isotropic porosity). There is therefore a link between granulometry, forcing and slope. This allows to define two characteristic cases: dissipative beach and reflective beach, see Fig. V.3. In reality, there is obviously a whole range of intermediate conditions.

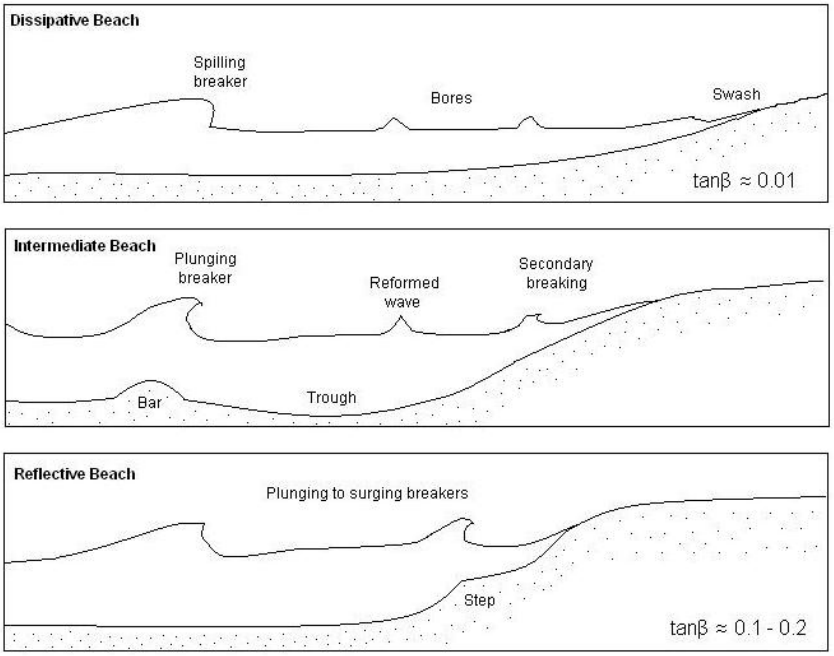


Figure V.3 – Beach classification taken from Wright and Short [215]. The classification depends on the angle of the beach slope β and on the wave condition.

Dissipative beaches are generally characterised by a fine sediment (sand or loamy sand) and gentle slopes. On these beaches, it is for excepted very small waves, offshore-generated waves break well before the shore (spilling waves). Throughout the surf zone, these waves dissipate a part of their energy

(friction, turbulence) and transfer it to longer waves, the so-called infragravity waves. The more dissipative the conditions are (gentle slope, large waves), the greater the dissipation and transfers are. Close to the shore, the sea state combines a variable amount of remaining post-breaking bore-shaped gravity waves and much longer infragravity waves, depending on the beach profile. An example is illustrated in Fig. V.4a. Within the beach, the velocities are about six orders of magnitude lower than the free surface velocities (1/10-1/100 mm/s), with mainly pressure waves that propagate. Saturation is highly variable and linked to sand properties (porosity) and local pressure field (capillary processes, formation of meniscuses, *etc.*). Then, large seepage face and capillary fringe may develop. The porous medium acts as a low-pass filter: the infragravity waves will have much more influence (deeper, stronger) than other waves.

Reflective beaches are generally characterized by coarser sediments, such as coarse sand or gravel, and steep slopes. There is little dissipation and energy transfer to the low frequencies (weak infragravity energy waves). The wave breaking occurs close to the shore, often forming plunging or surging breakers. It results in a thin tongue of turbulent water sheet whose tip is propagating rapidly with a lot of sediment transport. The phenomenon is analogous to a dam-break flow over a dry bed. An example is illustrated in Fig. V.4b. Due to much higher porosity than sandy beaches, there is a significant infiltration of water. Groundwater flows in coarse-grained beaches are fast, and the saturation is binary: either 100% or nothing.

(a) Seven Mile beach, Australia: a dissipative beach. Only small bores are reaching the swash zone. After a swash rundown, it results in seepage (glossy appearance) and capillary fringe (matte appearance)

(b) Avoca beach, Australia: a reflective beach. A bore is reaching the swash zone in the background and uprush occurs in the foreground. The swash zone is close to the shoreline

Figure V.4 – Examples of hydrodynamic processes in the swash zone.

V.1.2 Issues on groundwater beach dynamics

Beach groundwater dynamics has received a growing attention during the last two decades due to its importance in a number of physical, biological and chemical processes. The determination of groundwater transport through beaches or coastal barriers remains a key challenge in the prediction of salinity, contaminants and nutrients exchanges between the open sea, the lagoon and the aquifer which in turn control crucial biogeochemical cycles, see *e.g.* [216–222]. The beach groundwater circulation is mainly controlled by the combination of the cross-barrier gradients due to tides or large-scale fluctuations of the mean water level and the wave forcing at the shoreline, including swell and long

infragravity waves [4, 6, 7, 223, 224]. Furthermore, a smaller scale, the percolation flows in the swash zone have early been suspected to affect swash zone sediment transport and beach morphodynamics. The first effect is the loss of volume of the swash tongue, only active on coarse grained beach. The two other processes involved in sediment transport modification by in-exfiltration flows through the beachface are, on one hand, the modification of the boundary layer shear stress (the "ventilated bed" effect [225–228]) and, on the other hand, the change of the relative weight of sediment. Significant research efforts have been engaged to estimate and model the net effect of these processes on sediment dynamics [8, 9, 229–237]. Overall, while the impact is major for coarse-grained beaches, it is now generally acknowledged that the ex-infiltration effects on sediment transport for fine sand beach are overpowered by the other physical processes acting in the swash zone morphodynamics [4, 238]. This is in line with the mixed success met by beach drainage systems in trying to alter beach morphodynamics by controlling groundwater head [239, 240].

Owing to the difficulty to measure experimentally the groundwater parameters, and in particular to monitor their spatio-temporal dynamics over a relevant area, a series of numerical approaches have been proposed during the last few decades. First models based on Darcy's theory and purely horizontal groundwater dynamics (Dupuit-Forchheimer assumption) [223, 241] allowed to describe the tidal dynamics. Similar assumptions of purely horizontal flows were made in the coupled approach using Boussinesq model for waves propagation and Darcy-Forchheimer within the porous medium [232, 242]. Finer vertically-resolving model have been proposed based on the Darcy-related 2D Laplace law to model groundwater dynamics forced by tides and waves [8, 243] and further refined to describe in-exfiltration bed exchanges [237]. The groundwater circulation cell highlighted by coupled VOF-RANS model [234] has been later confirmed by laboratory [238] and field [6] experiments. In coarse-grained sand or gravel beaches, the Darcy theory is still used but extended beyond its laminar theoretical framework with the Darcy-Forchheimer approximation [244–246]. Quasi-3D model have been proposed, relying on an *a priori* parabolic profile of the non-hydrostatic groundwater pressure field [247, 248].

Despite a significant effort has been deployed during the last two decades in the understanding and the modelling of beach groundwater dynamics, most of the studies have been (i) based on Darcy-Forchheimer approximation, *i.e.* totally ignoring the partial saturation and capillarity effect which are essential in fine grained sand beaches [5] and (ii) mainly focused on horizontal dynamics, although detailed field measurements demonstrated the presence of vertical head gradients and related circulation cells in sandy beaches [6]. Aiming to tackle these two major challenges in the fine sand beach context, the present study aims to model wave-resolved variably-saturated groundwater flow based on Richards' equation. Generalized Richards' equation are generally used to model surface/subsurface flow in hydrological large-scale context (see *e.g.* [249, 250]), but very few attempts have been made to implement Richards' equation in wave-forced environments [96].

V.1.3 Beach response to a single swash event

Sedimentary beaches are considered as a porous medium in which there is a water table: the beach water table. Waves are covering and uncovering alternately the beach whose boundary is mobile, porous and characterized by complex flows of various natures which affect groundwater. That is why grasping groundwater dynamics as a whole in the swash zone is difficult. Laboratory experiments and field studies have revealed a typical circulation pattern during the swash cycle illustrated in Fig. V.5. The flow is mainly directed offshore, with inflow in the upper part of the swash zone and exfiltration in the lower part [4, 6, 238]. However, such features are expected to strongly depend on the saturation state of the beach and the duration and location of uprush/backwash phase, with significant consequences

on seepage flows or rise of water table [229, 251].

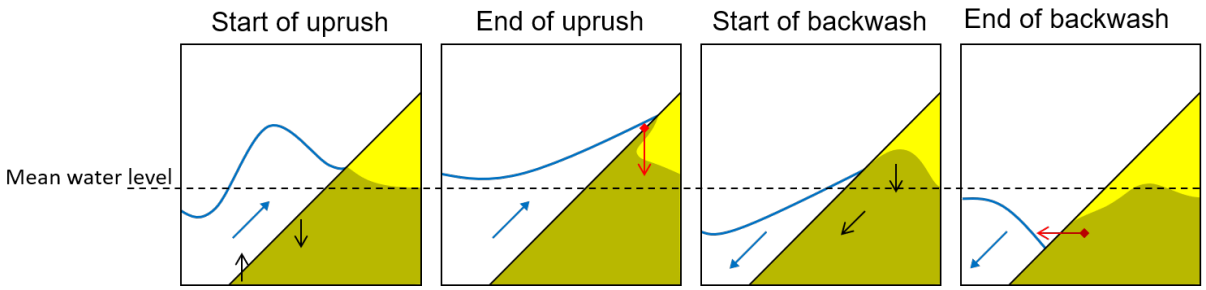


Figure V.5 – Conceptual view of the main flow patterns of beach groundwater induced by swash. Blue lines and arrows are for surface dynamics, black arrows for groundwater dynamics, shaded area depicts the saturated beach and red arrows for infiltration/seepage.

V.1.4 Coupling

Surface and subsurface flows are a unified system by nature. For technical reasons (space/time scales, computational solvability), they are often divided into two submodels which work together though interface conditions making a coupling. Numerically, different coupling schemes are possible, see the review of Furmann [252] or the work of Caviedes-Voullième *et al.* [96]:

- **forced coupling**, also known as, degenerated uncoupling, external coupling or weak coupling. It is a one-way coupling without feedback, from surface to subsurface flow. This type of coupling only requires (spatially and temporally variable) input data for interface conditions (boundary conditions);
- **parallel coupling**. Surface and subsurface models are solved separately with their own time step but with the interface conditions (source terms) evolving according to the updated results. This type of coupling presents a feedback but it is always shifted, in a kind of interlacing;
- **iterative coupling**, also known as alternating coupling or strong coupling. Solution of surface and subsurface models are harmonised at each time step via an interface condition with an iterative procedure. This coupling is the most complete but require additional algorithmic and computational efforts;
- **full coupling**. Solution of surface and subsurface models together with interface conditions are obtained simultaneously.

Moreover, theoretically, interface conditions can come from different ideas, for instance mass conservation principle or Beavers-Joseph-Saffman conditions.

The simplest one has been chosen for this study: the forced coupling. Based on the solution of surface water submodel, a boundary condition is prescribed for the groundwater model. It means submodels are solved in succession, and so, the surface submodel cannot be corrected by feedback. One benefit of forced coupling is that implementation is straightforward because no infiltration/exfiltration model is necessary. This technique can approximate sandy beaches where pore velocities are too slow to affect surface water (dissipative beach). Nevertheless, it goes without saying that coupling should gain in maturity by reckoning with an infiltration model or mass conservative principle, even without considering retroaction.

The non-hydrostatic phase-resolving SWASH surface wave model is used here to simulate propagation of free surface waves according to selected input parameters at the offshore boundary. It is used here in a two-layer configuration, all parameters being kept by default [253]. The space and time resolved bed pressure are extracted in order to force Richards' equation by a space and time dependant Dirichlet boundary condition for hydraulic head. Data are interpolated linearly from the space-time grid of SWASH to the space-time grid of *Rivage*. The swash tip is treated specifically by extrapolation to correspond to the closest discretization of *Rivage*. To complete this approach, a so-called seepage boundary condition [77] is prescribed wherever there are dry conditions, beyond the swash tip. In this way, the model can capture seepage faces, between surface water and a higher phreatic surface. The seaward seepage flow along beachface, as water is discharging, does not influence surface water. This forced coupling corresponds to a dynamic forcing boundary condition described in Section I.5.

V.2 Idealized beach

In the present study, an idealized long wave swash event is tested to evaluate the ability of the model to simulate groundwater response to swash cycles. To this end, a weak coupling is considered: wave-resolving bed pressure data provided by the open-source SWASH code [253] is used as a beachface boundary condition in the *Rivage* code. Qualitative comparisons are carried out between numerical results from the model and observations from laboratory and experimental measurements found in the literature [4, 6, 8, 237]. Distinctive features of the unconfined coastal aquifer during swash cycles are selected and commented. In particular, the analysis focuses on the evolution of seepage location across the beachface as well as global and local pattern of pressure and velocity field. The goal of this study is to present a first case handled by a promising numerical tool. The findings will assist future studies to quantify effects of swash on groundwater flow in sandy beaches for more realistic cases.

The beach considered is a straight slope of ratio 1:35 which spans over 105 meters. Figure V.6 shows the beach profile and the mesh used. Y-axis is multiplied by 20 to make it more visible.

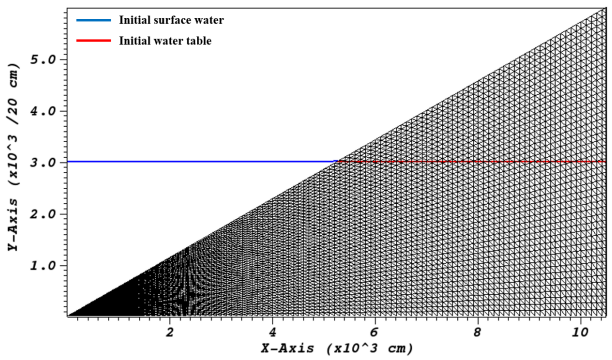


Figure V.6 – Topography of the idealized beach and mesh used for the simulation.

Hydraulic properties are taken from Vachaud's relations (see Tab. I.1) with the same set of parameters of Vauclin's experiment for water table recharge in a sandy box from Section IV.6.1: $A = 2.99 \times 10^6$, $B = 5.0$, $C = 40000$, $D = 2.90$, $K_s = 35$ cm/h, $\theta_s = 0.30$ and $\theta_r = 0$. Initial conditions are hydrostatic with water table level at the same height of surface water: $h_0 = 1.5$ m, see Fig. V.6. The bottom is an impervious boundary: $q \cdot n = 0$. The water table and head are kept fixed on the inland boundary: $h = 1.5$ m. The beachface is monitored by the dynamic forcing boundary condition with data from the SWASH code. In SWASH, a sinusoidal time-series of free surface elevation is imposed

at the offshore boundary to simulate a long infragravity-like wave of period 50 s and amplitude 0.3 m. These types of waves have been observed to be a major driver of beach groundwater dynamics of the Rousty microtidal dissipative beach [6]. This results in two swash events on the beachface. Simulated time lasts 120 s.

An overview of groundwater features during a swash cycle is presented. Then, results focus on uprush and backwash events. Some selected snapshots from the simulation show interesting observations for the idealized beach case. Surface water and water table are depicted by blue and white solid lines respectively. Hydraulic head contours are displayed together with velocities vectors.

A first observation concerns wave propagation. When a wave is moving, groundwater flows in the saturated zone respond immediately. Hydraulic head spreads vertically quite sharply under the wave front. Infiltration/exfiltration occurs at the back/front of the bore respectively. It is particularly visible in Fig. V.7 even if this phase is completely ideal because it is about the first wave moving through hydrostatic conditions.

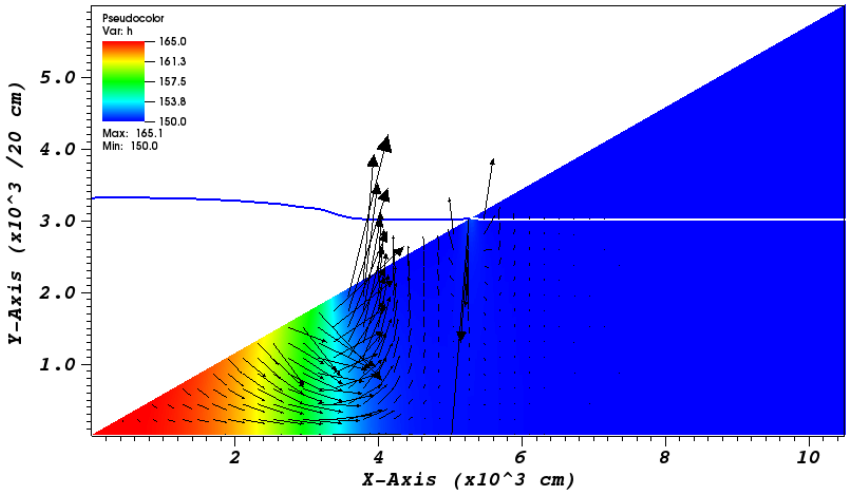


Figure V.7 – Snapshot of hydraulic head contours at $t = 12.5$ s during the first wave propagation on calm waters.

Throughout the simulation, a nearly hydrostatic state (no head gradients) is observed in the upper beach area. Moreover, flow divides into seaward and landward components near the uprush limit. Velocities magnitude is rather small everywhere, about 1 cm/h, except for uprush infiltration at the top where it reaches up to 160 cm/h. Around exit point of water table, a local overpressure is observed. It is probably caused by numerical artefacts due to the lack of mesh resolution for the seepage boundary condition. An overall pattern can be depicted for swash cycles. Two recirculation zones are attached respectively to the back and the front of the swash tongue. These two groundwater cells are interacting together and are evolving during uprush and backwash, growing or diminishing. Sometimes one vanishes completely, sometimes another one gets trapped by recharge induced by water table elevation.

During uprush, there is downward infiltration while the swash tongue is covering the beach. It is mainly concentrated in the upper zone, through the boundary in contact with the fully saturated part of the beach. Infiltration induces a rise of the water table near the exit point, with, as seen in Fig. V.8.

During backwash, a low hydraulic head zone emerges at the base of the swash zone where exfiltration occurs. Velocities are mainly seaward under the swash tongue moving back. Meanwhile, a small region

of high hydraulic head persists in the corner of the saturated zone where water table have been raised by the precedent uprush. It results in recharge and, as water is percolating inside the porous medium, flow divides between seaward and landward direction. Moreover, above the unsaturated zone which has been covered by uprush, water is infiltrating downward in a sharp wetting front. While the swash toe retreating to the sea, a seepage face is appearing which drops, locally but quickly, the water table what gives form to a small hump-like structure. Figure V.9 presents these observations.

The next incoming wave arrives just after the backwash. It pushes the region of low hydraulic head up. As this zone swept away, it reaches the remaining small region of high hydraulic head induced by the water table elevation and makes it disappear temporally. The incoming wave is bringing a new high hydraulic head front. This causes infiltration which arises water table again, ending seepage and erasing the hump-like structure of water table. Then, a new cycle begins. This intermediate state is shown in Fig. V.10.

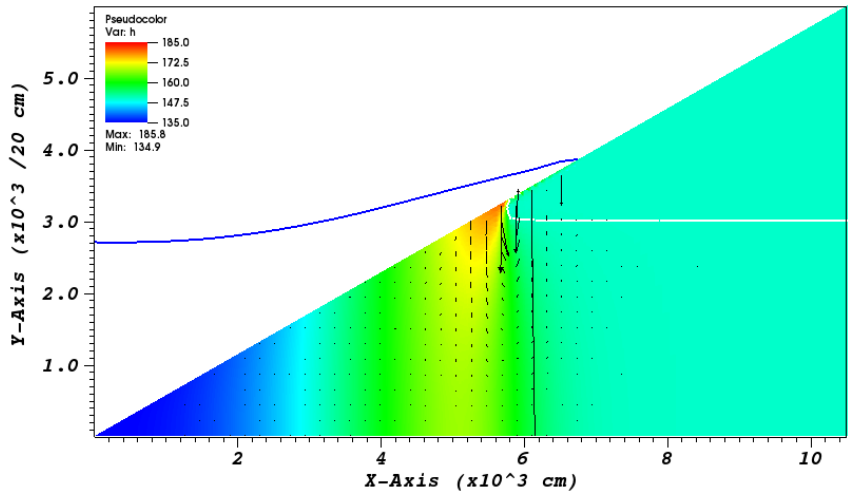


Figure V.8 – Snapshot of hydraulic head contours at $t = 37.5$ s for the end of the first uprush. Velocities range from about 0.3 to 0.6 cm/h in the beach. Maximum velocity is 35 cm/h for infiltration.

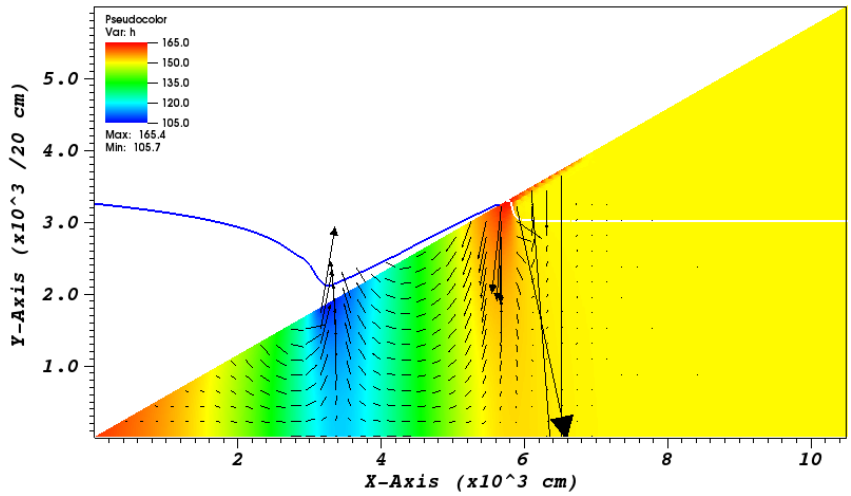


Figure V.9 – Snapshot of hydraulic head contours at $t = 57.5$ s for the end of the first backwash. Most velocities are less than 1 cm/h.

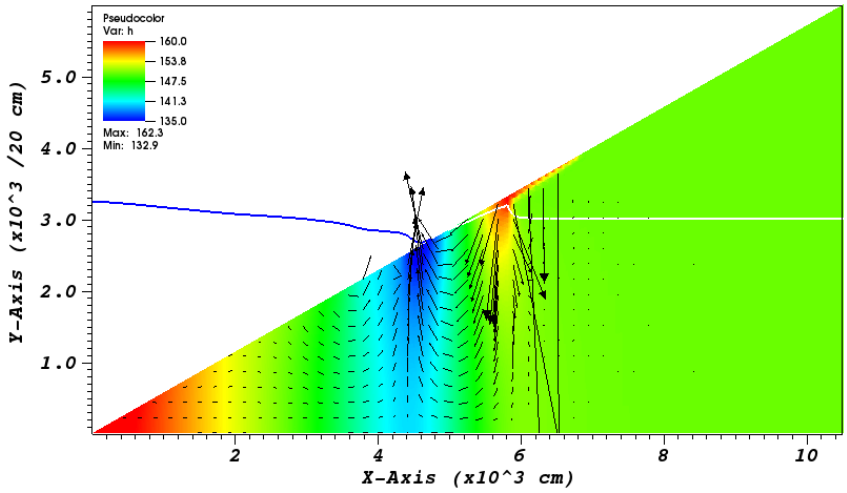


Figure V.10 – Snapshot of hydraulic head contours at $t = 67.5$ s for the next incoming uprush. Most velocities are less than 0.8 cm/h.

In the simulation, the offshore water level is similar to the inland water table, therefore no steady pressure gradient is observed across the beachface by contrast to field or laboratory studies where cross-beach gradients can be observed depending on boundary conditions [4, 6]. Within the beach, overall velocity magnitude is about 1 cm/h which is fairly comparable to measurements made in Sous *et al.* (2016). The only exception is for infiltration/exfiltration where maximum velocity order is about 100 cm/h. This seems to be less compared with the experiments of Turner *et al.* [4]. Further tests are now engaged in the exact configurations of the laboratory [4] and field [6] experiments to allow direct comparisons. As described by Perera *et al.* [237], the moving bore is carrying a sharp hydraulic front. Exfiltration occurs at its front and infiltration at its back when it meets the groundwater conditions of the precedent swash. A local superelevation of the water table has also been revealed in these observational studies due to the water table charge by short or IG wave forcing under the swash zone. Perera *et al.* showed a domed water table in their study [237]. The present benchmark case will be extended to a longer duration to test if the model is able to represent such feature. Water table elevation induces charge in this zone as long as the low-pressure zone does not affect it with the ongoing uprush. As a consequence, infiltration with flow division occurs near the uprush limit. Laboratory experiments and numerical simulations [4, 8] both predict this feature but Sous *et al.* do not observe it [6]. During the simulation, seepage face has small length because water is able to pour out from sand quite fast following surface water. As a consequence, a small hump-like structure appears during backwash because of seepage along the boundary. This last feature is not described in any study found, but further tests are ongoing with more realistic configurations and longer wave forcings to better understand involved processes. The two circulation zones with infiltration/exfiltration in the upper/lower part of the beach are also identified in the simulation from Li and Barry [8] and experiments from Turner *et al.* [4]. Infiltration and exfiltration involve greater pore velocities than in other parts of the beach. Like it is depicted by Perera *et al.* [237], gradients brought by the wave are much larger than the average values in the beach due to larger differences in water depth across the wave.

V.3 Swash groundwater dynamics prospects

The idealized beach case gives confidence, at least qualitatively and for order of magnitude, in the ability of *Rivage* model to treat swash groundwater dynamics in sandy beaches. Two additional cases

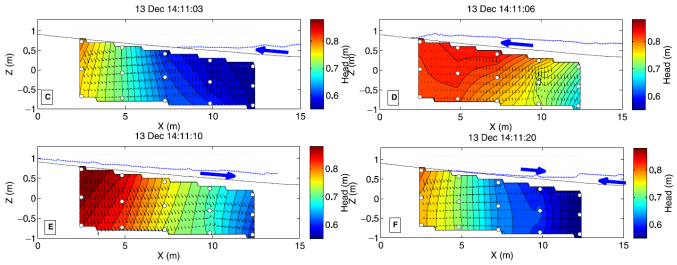
are considered. They are under construction and analysis.

The first case is provided by the laboratory swash experiments of [3]. A series of dambreak-generated bore-driven swash have been performed to study subsurface processes in the swash zone. The experimental set-up has been described in details in a series of studies, see *e.g.* [3, 254, 255]. The initial interest is mainly placed on coarse-grained beaches [245] but finer sediment is also tested [3]. This experiment remains a test-case of primary interest because it is one of the rare wave-resolved experiments on swash groundwater dynamics with detailed visualisation. Waves induce a mound in water table due to recharge. The development of corner wetting fronts above beach surface with sharp gradients in saturation and pressure head is a challenging issue for model validation.

The second case concerns the BARDEX II experiments. They were performed in July 2012 at the Delta Flume in the Netherlands [4, 235]. One of the main goals of this large scale experimental project was to characterize, under controlled laboratory conditions, the effect of groundwater flow on the hydro- and morphodynamics in a sand barrier. In particular, we focus here on a series of experiments dedicated to the effect of waves and cross-barrier mean water level gradients on the groundwater dynamics.

Finally, *Rivage* model will be used for in-situ experimental data collected during the ROUSTY campaign [6, 7], see Fig. V.11 for an example. It is field experiment on a microtidal beach in Camargue (France) called Rousty, see Fig. V.11. The study provides a comprehensive description of the groundwater dynamics under sandy beach swash zone thanks to complete measures about horizontal and vertical head gradients. A time-averaged and time-resolved analysis highlight respectively the groundwater circulation pattern under the swash zone and typical groundwater response to swash event. This experimental dataset is a challenging opportunity to test the *Rivage* model and gain new insights into swash groundwater dynamics.

(a) Photo taken after a storm showing an overview of the experimental set-up at the Rousty beach, France



(b) Snapshots of groundwater head contours and estimated velocity field for a swash event. Velocity magnitudes are between 10^{-4} and 10^{-6} m·s^{-1}

Figure V.11 – Examples of experimental data collected during the ROUSTY campaign. Taken from Sous *et al.* [6].

Summary

This chapter deals with a numerical model supported by the *Rivage* code and built around Richards' equation so that it can simulate wave-driven groundwater flows in sandy beaches. By generating waves, the SWASH code computes pressure on the bed to enforce Richards' equation with a Dirichlet boundary condition. For this forced coupling, groundwater does not hold sway over surface water but this influence can be neglected if pore velocity is slow compared to wave velocities as it happens for beaches with fine grains. What is more, exfiltration by seepage can be modelled through the seepage boundary condition prescribed on the beachface wherever there is no enforced pressure. Based on the simulation results from an idealized beach case, the following outlines can be drawn for a swash cycle:

- Groundwater pressure stays nearly hydrostatic under dry area, in the upper beach. Incoming waves propagate groundwater pressure front;
- The swash tongue causes overall infiltration during the bed flooding what elevates water table next to beachface and creates sharp wetting front in the unsaturated zone lapped by swash. For charge induced by water table, flow divides into landward and seaward components;
- During backwash, seaward groundwater can be observed under the swash tongue moving back to the sea. Water table exit point follows the swash toe with late because of sand hydraulic properties, resulting in seepage;
- Two localized groundwater circulation zones evolve and interact during the simulation. One, associated with low pressure and exfiltration, is attached to the swash base during backwash and to the forefront of the next wave during uprush. The other one, associated with high pressure and infiltration, is attached to water table recharge during backwash and to the swash wave during uprush.

This general scheme is consistent with observations made by experimental studies in the literature. Further investigations are needed to assess, on the one hand, numerical effects on the solution, such as discretization level or boundary conditions and, on the other, physical parameters such as constitutive laws, beach topography or incoming waves.

Conclusion

In this thesis, the main goal consisted in developing accurate, robust and efficient simulations of Richards' equation to model beach groundwater. For this purpose, a new strategy has been formulated for the solution of Richards' equation on the basis of DG method with BDF and AMR.

Summary

The Richards' equation approximation can reach high-order both in space and time, handle seepage, dynamic boundary conditions, heterogeneous porous media and deal with sharp drying/wetting fronts. Issues outlined in the introduction have been tackled and further discussed throughout the thesis so the following concluding remarks can be drawn:

- Richards' equation has been derived underlining clear hypotheses for the modelling;
- The seepage boundary condition has been directly incorporated into the DG weak formulation which makes its treatment natural and simple;
- The adaptive time stepping has allowed the nonlinear iterative solver to converge giving robustness;
- Mesh adaptation has been employed to monitor spatial errors of the Richards' equation by capturing the moving wetting fronts thanks to an estimation-based *a posteriori* error indicator;
- The AMR has been based on a block structure to surround large regions where wetting fronts move aiming to save computational time by avoiding systematic refinement. It also prepares a future parallelization;
- Oscillations in sharp wetting fronts of Richards' equation have been interpreted within the context of convection-diffusion equation. Wetting fronts have been considered as internal layers due to the nonlinear varying, and possibly vanishing, diffusivity as well as material heterogeneity and anisotropy. A local Péclet number has been exhibited to highlight troublesome regions. A WDG method has been used to allow for jumps in the solution which better approximate the sharp internal layers. Mesh adaptation and the weighted framework work in synergy to capture and resolve sharp wetting fronts through the proposed *a posteriori* error indicator.

Chapter I brought some insights about Richards' equation modelling possibilities. This nonlinear parabolic equation simulates an air-water system but considering only water phase. Thus, it is very useful to model flow in variably-saturated porous media but there are limitations such as air phase must stay continuously connected. Hydraulic properties suppose porous medium phases at equilibrium, disregarding hysteresis. Richards' equation modelling is still rich taking into account gravity and capillary effects. This double behaviour associated with nonlinearity make Richards' equation difficult to solve numerically for a wide range of applications.

Chapter II introduced the numerical methods used for Richards' equation. Space discretization was done by DG methods. DG weak formulation give flexibility to handle hybrid non-confirming unstructured mesh, local and arbitrary approximation order, special boundary conditions. Backward differentiation formula methods are used for time discretization to reach high-order and provide stability. Nonlinear solving lies on fixed-point iterations whose convergence is controlled by an heuristic adaptive time stepping. Some implementations choices in *Rivage* code were presented for solvers, basis functions or numerical integration.

Chapter III proposed the adaptive strategy based on AMR. A block-based structure saved computational time and prepares for future parallelization. Mesh adaptation is driven by error indicators supported by either problem physics or *a posteriori* estimation. They are combined with WDG framework to monitor jumps and flux in the solution. Wetting fronts may be so sharp that a discontinuity better approximate Richards' equation solution.

Chapter IV used several test-cases. Various 1D infiltration into soil column allowed to catch Richards' equation dynamics, to validate the code and to test capturing techniques by mesh adaptation and the weighted framework. In particular, jumps were introduced by the weighted penalty which prevents oscillation propagation when the front is sharp. The estimation-based error indicator revealed to perform very well to drive mesh adaptation cost-effectively. In addition, high-order time scheme were necessary to reach high-order accuracy when spatial errors are small. Adaptive time stepping was needed to make nonlinear convergence reliable. The code was able to simulate various dynamics (recharge/drainage) of laboratory experiments. The multi-materials La Verne dam wetting is a demanding case with experimental data. Comparison showed very good agreement with the system dynamics. Besides, it demonstrated the strategy ability to deal with sharp fronts from heterogeneity or nonlinear varying diffusivity. Mesh adaptation captured fronts with the estimation-based error indicator and the WDG method adjusted jumps size. Without providing a perfect solution, this strategy is very cost-effective compared to the mesh refinement required to improve the results.

Chapter V applied the method to groundwater in sandy beaches. The Richards' equation model was able to render main features observed experimentally even though only a forcing coupling is used. This is demonstrated by an idealized beach case where global and local space-time scales are modelled properly.

Perspectives

The strategy leads to mass conservative, cost-effective and reliable solution of Richards' equation. However, the method holds heuristic parameters so that questions remain open regarding the numerical analysis, the *hp*-decision making, the nonlinear convergence or the error balance. All these aspects have great potential of improvements and will be investigated in future studies to optimize the solving of Richards' equation.

The following suggestions are provided for future research:

- A work of validation and confrontation is under way with cases involving laboratory and largescale experiments for groundwater in sandy beaches. An application to a field experiment on Rousty beach is aimed to provide new insights about swash groundwater;
- Volume averaging method would be interesting to continue in order to derive clear assumptions about flow regime and porous medium. First, it will contribute to the field of volume averaging by proposing dimensional analysis based on different characteristics lengths. Secondly, it will precise the range of validity for the fundamental relation modelling porous medium flux which

is then used in Richards' equation. Finally, volume averaging is also a powerful tool to elaborate coupling models between subsurface and surface flows. This will contribute building a coupling more elaborate than a forcing;

- Penalty parameters provide stability for DG methods and enforce boundary conditions. However, they can also degrade numerical properties with arbitrary values, particularly condition number, leading to instabilities. Besides, they offer a way to incorporate into the scheme some partial knowledge about the solution. Numerical analysis and further tests would be carried on to derive a definition convenient for the problem or, at least, suitable bounds;
- Adaptive time stepping controls nonlinear iterative convergence. However, the method does not prevent the time step to decrease so much that computation is not feasible. Besides, the method is heuristic and it is difficult to know if computational time is saved since swing factor is a user-defined parameter. One possibility would be to use a time error estimator. Another possibility would be to study convergence of the nonlinear procedure which depends on the time step and to deduce a method. Local time stepping might be an idea to save computational time but it is difficult to make with implicit schemes;
- Time order selection is fixed for now in the code. This part should be fully automatic providing the best BDF method to guarantee stability. This is usually done by the mean of error estimators or stability analysis;
- *hp*-Adaptation would be a profitable for Richards' equation solution by performing mesh refinement or order enrichment following the one is better. Lots of techniques are already available in literature to do *hp*-decision making. Some remarks in this thesis make think that subparts of the estimation-based error indicator may drive the decision between *h*- and *p*-adaptation. Even though this *hp*-adaptation will not be optimal, the technique would be very simple and cost-effective. Further tests are needed and a true *a posteriori* error estimation analysis could help;
- Adaptivity is an issue for the refinement frequency and the selection of elements to refine. Automatic thresholding based on a distribution function would be a possibility to mark elements for refinement/coarsening. Deriving velocity of wetting fronts would be a way to define refinement frequency given the block size;
- Parallelization and 3D are natural extensions for *Rivage* code to handle more complex cases and save computational time. Even though the code is prepared for this (block-based structure leading to domain decomposition and general implementation), it will be not straightforward since 3D involves lots of element geometries and time scheme is implicit. Besides, adding a module to treat hyperbolic equation would be a possibility to investigate in order to incorporate surface flow models and make a proper coupling.

Valorizations

Several parts of the thesis were presented during conferences and seminars whose some of them provided proceedings. One article has been proposed to a journal for publication where the adaptive DG strategy is described to get Richards' equation solution. Another one is planned to introduce and discuss groundwater applications to sandy beaches. These works are listed in List of Publications, page 145.

Appendices

Appendix A - Spatial averaging theorem

The derivation of the spatial averaging theorem may use different routes. The one presented here is done in the framework of convolution. The demonstration involves change of variables. That is why derivatives operators are indexed by the variable at stake.

A filtering operator is defined as the convolution product between a mask function m and the considered intrinsic quantity $\xi_{\mathscr{F}} = \mathbb{1}_{\mathscr{F}}\xi$. This is called phase average:

$$\begin{aligned} \langle\xi\rangle(x) &= m * \xi_{\mathscr{F}}\Big|_x \\ &= m * (\mathbb{1}_{\mathscr{F}}\xi)\Big|_x \\ &= \int_{\mathbb{R}^3} m(-y)\mathbb{1}_{\mathscr{F}}(x+y)\xi(x+y)\,\mathrm{d}y. \end{aligned} \tag{V.1}$$

The mask function m defines a weighted window which acts as a kernel in volume averaging. An analogy is possible with the theory of image processing. So m is set as:

$$m(y) = \begin{cases} \dfrac{1}{|V|} & \text{if } |y| \leq r_0, \\ 0 & \text{otherwise.} \end{cases} \tag{V.2}$$

Since m is an even function, it writes:

$$\langle\xi\rangle(x) = \int_{\mathbb{R}^3} m(y)\mathbb{1}_{\mathscr{F}}(x+y)\xi(x+y)\,\mathrm{d}y. \tag{V.3}$$

The average of a constant quantity ξ_0 should be ξ_0. Yet, it stands:

$$\langle\xi_0\rangle(x) = m * (\mathbb{1}_{\mathscr{F}}\xi_0)\Big|_x = \xi_0 m * (\mathbb{1}_{\mathscr{F}})\Big|_x = \xi_0\langle 1\rangle(x) \neq \xi_0. \tag{V.4}$$

That is the reason why a second filtering operator is needed. It is called intrinsic average:

$$\overline{\xi}(x) = \frac{\langle\xi\rangle(x)}{\langle 1\rangle(x)} = \frac{\langle\xi\rangle(x)}{\Phi(x)}. \tag{V.5}$$

It is interesting that the intrinsic average is idempotent, that is to say the intrinsic average of the intrinsic average is equal to the intrinsic average. However:

$$\overline{\overline{\xi}} = \overline{\overline{\xi} + \widetilde{\xi}} = \overline{\overline{\xi}} + \widetilde{\overline{\xi}}. \tag{V.6}$$

In order to have $\overline{\overline{\xi}} = \overline{\xi}$, it is necessary that $\overline{\widetilde{\xi}} = 0$. In other words, the intrinsic average of the spatial deviation is zero. However, to be in accordance with the idempotency property, there are:

$$\begin{aligned}
\Phi(x)\overline{\overline{\xi}}(x) &= \langle\xi\rangle(x) \\
&= \int_{\mathbb{R}} m(y)\mathbb{1}_{\mathscr{F}}(x+y)\overline{\xi}(x+y)\,\mathrm{d}y \\
&= \int_{\mathbb{R}} m(y)\mathbb{1}_{\mathscr{F}}(x+y)\overline{\xi}(x)\,\mathrm{d}y + \int_{\mathbb{R}} m(y)\mathbb{1}_{\mathscr{F}}(x+y)\mathsf{T}\nabla\overline{\xi}(x)y\,\mathrm{d}y \\
&+ \int_{\mathbb{R}} m(y)\mathbb{1}_{\mathscr{F}}(x+y)\frac{1}{2}\left(\mathsf{T}yH_{\overline{\xi}}(x)y\right)\,\mathrm{d}y + \mathcal{O}\left(\mathsf{T}yy\right) \\
&= \underbrace{\langle 1\rangle(x)}_{\substack{\text{Order 0}\\ \text{moment}}}\overline{\xi}(x) + \underbrace{\langle y\rangle(x)}_{\substack{\text{Order 1}\\ \text{moment}}}\nabla\overline{\xi}(x) + \frac{1}{2}\underbrace{\langle\mathsf{T}yy\rangle(x)}_{\substack{\text{Order 2}\\ \text{moment}}}H_{\xi}(x) + \mathcal{O}\left(\mathsf{T}yy\right) \\
&= \Phi(x)\overline{\xi}(x) + \underbrace{\langle y\rangle\nabla\overline{\xi}(x) + \frac{1}{2}\langle\mathsf{T}yy\rangle H_{\xi}(x) + \mathcal{O}\left(\mathsf{T}yy\right)}_{=-\Phi(x)\overline{\widetilde{\xi}}=0}. \qquad \text{(V.7)}
\end{aligned}$$

A dimensional analysis gives:

$$\langle y\rangle(x)\nabla\overline{\xi}(x) = \mathcal{O}\left(\frac{r_0\Phi(x)\overline{\xi}}{L}\right), \qquad \text{(V.8)}$$

$$\langle\mathsf{T}yy\rangle(x)H_{\xi}(x) = \mathcal{O}\left(\frac{r_0\Phi(x)\overline{\xi}}{L^2}\right). \qquad \text{(V.9)}$$

So the intrinsic average is idempotent provided that $\dfrac{r_0}{L} \ll 1$.

Demonstration of spatial averaging theorem:

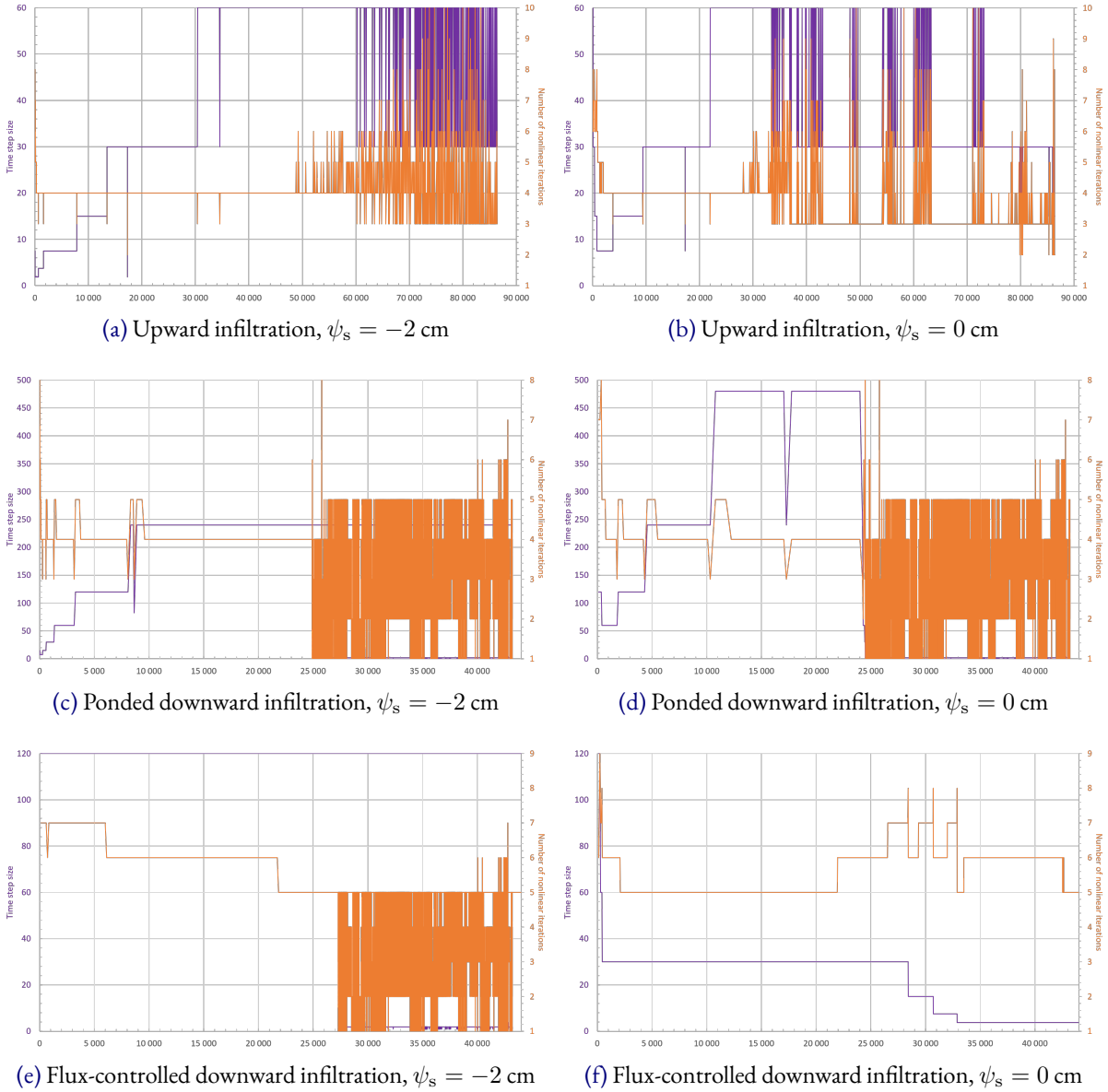
$$\begin{aligned}
\nabla_x\langle\xi\rangle\Big|_x &= \nabla_x\left[\int_{\mathbb{R}^3} m(y)\mathbb{1}_{\mathscr{F}}(x+y)\xi(x+y)\,\mathrm{d}y\right] && \text{(V.10)}\\
&= \nabla_x\left[\int_{\mathbb{R}^3} m(z-x)\mathbb{1}_{\mathscr{F}}(z)\xi(z)\,\mathrm{d}z\right] && \text{(V.11)}\\
&= \int_{\mathbb{R}^3} \nabla_x[m(z-x)]ind_{\mathscr{F}}(z)\xi(z)\,\mathrm{d}z && \text{(V.12)}\\
&= -\int_{\mathbb{R}^3} \nabla_z[m(z-x)]\mathbb{1}_{\mathscr{F}}(z)\xi(z)\,\mathrm{d}z && \text{(V.13)}\\
&= -\int_{\mathbb{R}^3} (\nabla_z[m(z-x)\mathbb{1}_{\mathscr{F}}(z)\xi(z)] - m(z-x)\nabla_z[\mathbb{1}_{\mathscr{F}}(z)\xi(z)])\,\mathrm{d}z && \text{(V.14)}\\
&= \int_{\mathbb{R}^3} m(z-x)\nabla_z[\mathbb{1}_{\mathscr{F}}(z)\xi(z)]\,\mathrm{d}z - \int m(z-x)\mathbb{1}_{\mathscr{F}}(z)\xi(z)\boldsymbol{n}\,\mathrm{d}\sigma && \text{(V.15)}\\
&= \int_{\mathbb{R}^3} m(y)\nabla_z[\mathbb{1}_{\mathscr{F}}(x+y)\xi(x+y)]\,\mathrm{d}y - \int m(z-x)\mathbb{1}_{\mathscr{F}}(z)\xi(z)\boldsymbol{n}\,\mathrm{d}\sigma && \text{(V.16)}\\
&= \langle\nabla_x\xi\rangle\Big|_x - \int m(z-x)\mathbb{1}_{\mathscr{F}}(z)\xi(z)\boldsymbol{n}\,\mathrm{d}\sigma && \text{(V.17)}
\end{aligned}$$

And thus:

$$\langle\nabla_x\xi\rangle\Big|_x = \nabla_x\langle\xi\rangle\Big|_x + \int m(z-x)\mathbb{1}_{\mathscr{F}}(z)\xi(z)\boldsymbol{n}\,\mathrm{d}\sigma \qquad \text{(V.18)}$$

Appendix B - Complementary results

Additional numerical data from Vogel-Van Genuchten-Cislerova test-cases (see Section IV.4) are provided to show the interdependence between the number of nonlinear iterations and the time stepsize. User-defined tolerances for stopping criteria are $\epsilon_1 = \epsilon_2 = 10^{-4}$ and nonlinear parameters for time stepping are $m_{\text{it}} = 3$, $M_{\text{it}} = 7$, $W_{\text{it}} = 10$, $\lambda_{\text{amp}} = 2$ and $\lambda_{\text{red}} = 0.5$. The results are gathered and presented in Fig. V.12. Small time stepsizes mean that convergence is difficult: it is the case for downward infiltrations, especially for the flux-controlled one. When the number of nonlinear iterations is oscillating rapidly, it means that convergence is difficult (solver failure for downward infiltration) or that amplification/reduction parameters are too strong to provide smooth nonlinear iterative process running over time (upward infiltration).



(a) Upward infiltration, $\psi_{\text{s}} = -2$ cm

(b) Upward infiltration, $\psi_{\text{s}} = 0$ cm

(c) Ponded downward infiltration, $\psi_{\text{s}} = -2$ cm

(d) Ponded downward infiltration, $\psi_{\text{s}} = 0$ cm

(e) Flux-controlled downward infiltration, $\psi_{\text{s}} = -2$ cm

(f) Flux-controlled downward infiltration, $\psi_{\text{s}} = 0$ cm

Figure V.12 – Complementary results for Vogel-Van Genuchten-Cislerova test-cases.

Appendix C - Three Minute Thesis – Ma thèse en 180 secondes

Ma thèse en 180 secondes (MT180) is an annual popular science competition accessible to French speaking PhD students worldwide. The concept is inspired from *Three minute thesis* (3MT) created by the University of Queensland in Australia. Participants present their research subject in French to a general and diversified audience with no background in the research area. The challenge is each student has three minutes exactly to give an intelligible and convincing talk with only one slide illustration. Competition is held in each university. Then, selected candidates competes each other successively at regional, national and international level.

I participated in MT180 Édition 2019. Université de Toulon organized a night show called *La recherche fait son show* on the 7th of March 2019 at Liberté theater in Toulon to decide between its candidates. My talk is available in video on the following link (in French): http://recherche-show.univ-tln.fr/saison-2-resultats-et-videos/. The corresponding speech is transcribed below with the slide illustration presented in Fig. V.13. An approximate translation of the speech in English is proposed hereafter.

« Ce soir, j'aimerais faire appel à l'enfant qui sommeille en vous. Ce pitchoun, haut comme trois pommes, qui fait ses pâtés de sable sur la plage. Imaginez la scène : tranquillement allongés à l'ombre d'un cocotier, vous écoutez la rumeur des vagues. Le rêve. Mais je vous rappelle que vous êtes retombés en enfance... Vous n'êtes pas là pour bronzer ! Non. Vous avez décidé d'aller, ici, sur le bord de plage pour construire votre château de sable. Vous entreprenez de creuser un trou et là... Ah ! C'est le drame ! De l'eau surgit du sable. Eh bien oui, il y de l'eau dans le sable de la plage. Ça, tout le monde le sait. Mais attendez... Que fait réellement toute cette eau dans le sable de la plage ? Mystère et boule de gomme. Je vais vous le dire. Chaque vague, par ses flots montants – infiltration – et ses flots descendants – exfiltration – crée une circulation souterraine, lente, tellement lente que les expériences datent d'il n'y a pas si longtemps. On parle de centimètres par heure. Et maintenant ? Maintenant, on veut en savoir plus. Est-ce que cela diffère selon le type de vagues, de sable ? Quelle est l'influence sur l'érosion ?
Et c'est là que ma thèse intervient. Je m'intéresse à la simulation numérique. Mesdames et messieurs, bienvenue dans le monde merveilleux du numérique ! Attention, pas celui d'Internet ou de votre tablette. Non, celui des mathématiques appliquées. Cela vous effraie ? Il ne faut pas car il y a un peu de magie là-dessous. Je vous avez prévenu, nous restons en enfance. Pour faire une bonne simulation numérique, il nous faut réunir trois ingrédients. D'abord, un modèle, une formule magique. Comme vedette de ma thèse, j'ai retenu l'équation de Richards qui décrit spécifiquement le mouvement de l'eau dans le sable même si de l'air est présent. Gare aux erreurs de modélisation ! Ensuite, j'ai besoin d'une incantation, une méthode numérique qui transforme mon équation pour la résoudre sur un ordinateur. Après consultation des grimoires, j'ai jeté mon dévolu sur une méthode assez originale qui demande des sacrifices – informatiques – en temps et mémoire mais qui par sa grande flexibilité permet d'atteindre une précision remarquable. Enfin, il faut préparer la potion dans un chaudron bien particulier, un programme informatique appelé code de calcul.
L'ensemble du procédé est délicat car je dois m'assurer que les calculs se font en un temps raisonnable et que la métamorphose ne dégrade pas les propriétés de l'équation, sans quoi j'obtiendrai des résultats faux. Or, si je suis capable de reproduire ce qui est observé, je pourrais simuler d'autres situations, le futur, mieux prédire l'érosion du littoral. Et c'est ça la magie des mathématiques. Face au réchauffement climatique et à l'urbanisation côtière, nos plages disparaissent. Nous y sommes tous attachés comme un enfant à son château de sable alors il serait temps d'en prendre soin. »

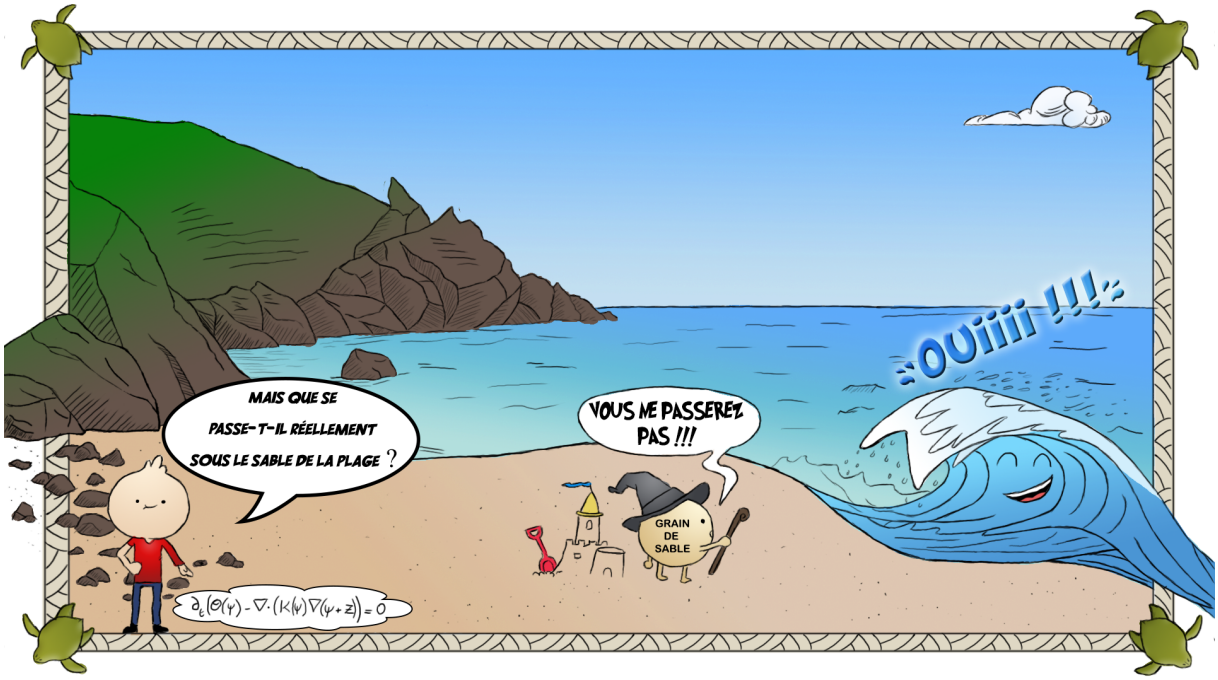


Figure V.13 – Illustration of the slide during the talk for MT180. "But what is really happening under beach sand?", "You shall not pass!!! (sand grain)", "Yeeees!!!".

"Tonight, I would like to call upon the inner child in you. This inner child, as tall as three apples, who makes sandcastles on the beach. Imagine the scene: quietly lying in the shade of a coconut tree, you listen to the rumble of the waves. The dream. But I must remind you that you fell back into childhood... You're not here to sunbathe! No. You have decided to go, here, on the beach, to build your sandcastle. You're going to dig a hole and then... Ah! It's a tragedy! Water rises from the sand. Well, yes, there is water in the sand of the beach. Everybody knows that. But wait... What is all this water really doing in the sand of the beach? That is a riddle wrapped in a mystery inside an enigma. I'm going to tell you. Each wave, through its rising flow – infiltration – and its descending flow – exfiltration – creates a slow underground circulation, so slow that experiments date back to not so long ago. We're talking about centimetres per hour. And now? Now, we want to know more. Does it depend on the type of waves or sand? What is the influence on erosion?

And that's where my thesis comes in. I am interested in computer simulation. Ladies and gentlemen, welcome to the wonderful world of numerical data! Be careful, not the digital on the Internet or the digital on your tablet. No, it's applied mathematics. Does that scare you? Don't be scared, because there's a little magic down there. I warned you, we stayed in childhood. To make a good numerical simulation, we need to combine three ingredients. First, a model, a magic formula. The star of my thesis, I chose Richards' equation, which specifically describes the movement of water in sand even though air is present. Beware of modelling errors! Then, I need a spell, a numerical method that transforms my equation to solve it on a computer. After consulting the books of spells, I set my sights on a rather original method that requires sacrifices — computing ones – in time and memory, but thanks to its greater flexibility allows to achieve remarkable precision. Finally, you have to prepare the potion in a very special cauldron, a computer program called a calculation code.

The whole process is tricky because I have to make sure that computation is done in a reasonable time and that the metamorphosis does not degrade the properties of the equation, otherwise I will get false results. However, if I am able to reproduce what is observed, I could simulate other situations, the future, better predict coastal erosion. And that's the magic of mathematics. Because of global

warming and coastal urbanization, our beaches are disappearing. We are all attached to beaches like a child with his sandcastle, so it is time to take care of it."

Résumé étendu en français

Présentation générale

La prédiction des écoulements dans les milieux poreux variablement saturés est un problème majeur pour de nombreux domaines en science et en ingénierie. Ces écoulements apparaissent par exemple dans la physique des sols et de l'environnement, l'hydrogéologie, l'agriculture ou l'industrie pétrolière pour des problèmes comme le transport de contaminants, les réservoirs de pétrole, les ressources en eau, les affleurements de nappe, *etc*. Les études expérimentales sont un outil fondamental pour décrire de tels processus et obtenir une bonne compréhension. Cependant, les études expérimentales ne sont pas toujours possibles, ni réalisables. La modélisation mathématique et la simulation numérique offrent une autre possibilité pour explorer et acquérir de nouvelles connaissances sur ces systèmes physiques.

La compréhension de la dynamique des écoulements souterrains dans les zones côtières est d'un intérêt majeur dans les domaines du génie côtier et du développement durable. Dans le présent travail, une attention particulière sera apportée aux applications des écoulements souterrains littoraux. Une meilleure compréhension de la circulation de ces écoulements est d'une importance primordiale pour analyser et prévoir un ensemble de processus physiques et biogéochimiques des zones littorales, tels que le transport des sédiments (stabilité du lit), la diffusion de matières dissoutes telles que les polluants ou les nutriments ou encore le mélange entre les eaux continentales (douces) et marines (salées). Chacune de ces questions requiert une attention particulière dans le contexte du changement climatique et de l'érosion [1, 2] (élévation du niveau de la mer, évènements de submersion, salinisation des terres arables) et de la pression anthropique croissante (urbanisation côtière).

Des études récentes en laboratoire [3, 4] et sur le terrain [5, 6] ont permis d'acquérir une bonne connaissance expérimentale des processus physiques régissant la dynamique des écoulements souterrains dans les plages sbleuses exposées à l'action des vagues. Les données récemment obtenues au cours de la campagne ROUSTY2014 permettront d'approfondir l'analyse de cette dynamique souterraine à plus grande échelle [7]. Malgré leur importance primordiale, les mesures expérimentales restent très coûteuses et sont limitées en termes de résolution et d'étendue temporelle et spatiale. Un défi scientifique majeur consiste à développer des modèles numériques robustes et efficaces pour mieux anticiper, via la réalisation de scénarios prédictifs, l'évolution de la zone littorale. Peu de modèles ont été présentés jusqu'à présent et leur champ d'application reste limité, voir *e.g.* [8, 9]. De nouveaux développements numériques sont nécessaires pour aborder les questions clés des écoulements souterrains de la plage induits par les vagues.

De nombreux modèles d'écoulements en milieux poreux ont été proposés et développés pour simuler les flux et processus de transport à différentes échelles. Les problèmes d'écoulement multiphasique présentent un comportement plus complexe que les problèmes d'écoulement monophasique car la dynamique des phases s'influence mutuellement. Dans ce contexte, le modèle d'écoulement multi-

phasique peut conduire à des difficultés de résolution ou avoir un niveau de détails trop complexe pour la gamme d'applications visées. Un modèle simplifié tenant compte d'hypothèses raisonnables constitue une alternative très intéressante.

Dans le présent travail, les écoulements dans des milieux poreux à saturation variable sont décrits par l'équation de Richards. Ce qui rend l'équation de Richards très attrayante est qu'elle modélise le milieu poreux dans son ensemble, y compris les zones saturées et non-saturées. Bien qu'elle néglige la phase d'air, l'équation de Richards tient compte des actions de la gravité et de la capillarité et peut donc modéliser des processus variés et complexes à différentes échelles. L'équation de Richards est une équation parabolique non-linéaire qui peut dégénérer en une équation elliptique dans des conditions de saturation complète. Son histoire commence avec la loi de Darcy formulée expérimentalement par Darcy en 1856 [10] pour les milieux poreux saturés. Puis, ce résultat a été étendu aux écoulements multiphasiques par Buckingham en 1907 [11] pour donner la loi de Darcy-Buckingham, qui sert de pierre angulaire pour la dérivation de l'équation de Richards. Cette dernière a été établie pour la première fois par Richardson en 1922 [12] mais a été attribuée uniquement à Richards qui a publié indépendamment l'équation en 1931 [13]. Les premières tentatives de résolution numérique remontent à la fin des années 1960 avec Rubin [14] ou Cooley [15] par exemple. Puis, à partir des années 1980, l'équation de Richards a été étudiée de manière approfondie, tant du point de vue théorique que numérique. Des résultats d'existence, de régularité et d'unicité ont été trouvés [16] et la solution a été évaluée par des simulations impliquant différentes méthodes numériques. Un des points d'intérêt concernait le solveur non-linéaire [17, 18]. Depuis les années 2000, les travaux portent principalement sur la stabilité de l'équation de Richards et le développement de simulations couplées et générales [19]. Ces dernières années, les recherches portent sur l'équation de Richards avec hystérésis et discrétisations spatiales/-temporelles avancées [20, 21] avec des approximations adaptatives en taille et/ou en ordre.

L'équation de Richards a été largement utilisée pour des simulations numériques par la communauté d'hydrogéologie, voir *e.g.* [19, 22]. Malgré cet effort de recherche, de nombreux défis numériques restent à relever pour résoudre l'équation de Richards, certaines simulations étant encore peu fiables et/ou coûteuses [23, 24]. Cette complexité numérique a empêché une utilisation plus générale du modèle de Richards pour un certain nombre d'applications. En effet, la solution de l'équation de Richards implique des fronts de saturation raides difficiles à résoudre car ils évoluent à la fois dans l'espace et le temps. En outre, la simulation doit traiter simultanément des régions non-saturées/saturées de nature parabolique/elliptique, une forte hétérogénéité et anisotropie du milieu poreux et éventuellement des conditions aux limites dynamiques. Ces aspects peuvent également conduire à de forts gradients agissant comme des discontinuités. Compte tenu des fonctions non-linéaires pour les propriétés hydrauliques, il est rarement possible de prévoir le comportement de l'équation de Richards et il peut être difficile d'obtenir des solutions numériques. En particulier, les schémas numériques ne convergent pas systématiquement, ce qui nécessite un traitement numérique minutieux, ou une discrétisation fine, rendant le calcul coûteux.

L'objectif de cette thèse est de développer des outils conduisant à une simulation de l'équation de Richards qui est :

- précise, dans le sens que la solution numérique doit être physiquement acceptable et converger vers une solution *exacte* lorsque les pas de temps et/ou la résolution du maillage sont raffinés ;
- robuste, dans le sens que la méthode doit fournir une solution quels que soient les paramètres physiques et numériques du problème ;
- efficace, dans le sens que le calcul doit être rentable en termes de temps de calcul et de ressources mémoire ;

- adapté aux écoulements souterrains de la plage induits par les vagues.

Le présent travail de recherche numérique vise la simulation des écoulements souterrains dans les plages sableuses comme principale application. Luijendijk *et al.* ont constaté que plus d'un tiers des côtes sont sableuses. L'analyse des données satellitaires sur une période de 33 ans (1984-2016) indique qu'un quart des plages de sable s'érodent à des taux supérieurs à 0,5 m/an. Les littoraux sableux présentent des avantages socio-économiques liés au tourisme et aux services écosystémiques. Ils font l'objet de vives préoccupations en raison du changement climatique, car l'élévation du niveau de la mer menace les plages sableuses d'une érosion, comme affirmé par Vousdoukas *et al.*. Cooper *et al.* [25] ont nuancé la disparition des plages sableuses qui peuvent se retirer à l'intérieur des terres. Les plages sableuses restent soumises à d'importants changements et peuvent disparaître si l'urbanisation bloque leur migration. En outre, du point de vue de la modélisation, les plages sableuses sont généralement des plages dissipatives à faible pente pour lesquelles les écoulements souterrains sont connus pour être lents avec une faible rétroaction sur les vagues incidentes. C'est pourquoi un couplage faible de l'écoulement de surface sur l'écoulement souterrain peut être envisagé pour cibler des simulations applicatives. Ces raisons font que la thèse priorise l'étude des écoulements souterrains pour les plages sableuses.

À cette fin, une méthode de Galerkine discontinue (GD) est choisie pour résoudre l'équation de Richards. Les méthodes GD sont basées sur une formulation variationnelle élément par élément, partageant des avantages des méthodes par éléments finis et par volumes finis. En particulier, elles sont localement conservatives, ce qui est crucial en mécanique des fluides [26]. De plus, la nature de la formulation GD permet de travailler sur des maillages non-conformes et de modifier localement le degré d'approximation polynomiale. C'est un avantage important car le Raffinement du Maillage Adaptatif (RMA), appelé adaptation *h*, et la des méthodes d'ordre élevé, appelée adaptation *p*, deviennent possibles [27]. Cel favorise une utilisation croissante des méthodes de GD pour les phénomènes de transport dans les milieux poreux tels que les problèmes d'écoulement diphasique [28, 29].

En outre, l'équation de Richards est connue pour être une équation différentielle raide avec une convergence difficile [17, 18, 30]. La discrétisation temporelle et le solveur non-linéaire doivent donc être traités avec soin.

En gardant à l'esprit que la méthode numérique doit fournir des résultats précis et robustes, mais efficaces en termes de temps de calcul, cette thèse constitue une opportunité de contribuer à l'avancement des méthodes GD adaptatives d'ordre élevé pour la simulation de l'équation de Richards. Ainsi, les réalisations suivantes sont attendues :

- L'équation de Richards est discrétisée par des méthodes GD primales qui conduisent à une formulation conservative, d'ordre élevé et un RAM pour obtenir une solution précise. La pénalisation impose faiblement les conditions aux limites dans la formulation GD pour la robustesse. Les formules de différenciation à rebours sont utilisées pour la discrétisation temporelle car elles sont connues pour être d'ordre élevé et stables pour les équations rigides ;
- Une méthode à point fixe associée à un échelonnement adaptatif du temps rend les itérations non-linéaires plus fiables ;
- Un indicateur d'erreur *a posteriori* est formulé pour l'équation de Richards approchée par des méthodes de GD. Cette estimation est décomposée en différentes contributions qui guident l'adaptation du maillage ;
- L'algorithme utilise une RMA par blocs pour gagner du temps de calcul.

Cette thèse propose une nouvelle stratégie pour étudier les simulations adaptatives en temps et espace pour l'équation de Richards. Le RMA sera utilisé pour capturer les fronts de saturation en mouvement

grâce à une estimation *a posteriori*. L'adaptation *h* est employée pour les écoulements diphasiques en milieux poreux [28] mais aussi pour l'équation de Richards [20, 31]. Tirant le meilleur partie de la flexibilité des méthodes GD, le RMA est combiné avec un cadre de Galerkine discontinu pondéré (GDP) qui permet une discontinuité dans la solution en fonction de la diffusivité non-linéaire. Cette approche a été formulée par Ern, Di Pietro et d'autres collaborateurs [32, 33] et par Proft et Rivière [34, 35] en suivant à des travaux antérieurs. L'application d'une telle stratégie à l'équation de Richards, combinant le RMA, l'estimation *a posteriori* et le cadre GDP, est la principale nouveauté de l'étude.

Un code dédié appelé *Rivage* a été conçu, développé et validé au cours de cette thèse. *Rivage* servira au sein du laboratoire pour l'étude plus approfondie de l'équation de Richards mais il sera également étendu à d'autres problèmes. Des cas supplémentaires basés sur des expériences en laboratoire et à grande échelle sont en cours de construction et d'analyse.

La thèse est organisée comme suit.

Le Chapitre I présente l'équation de Richards de manière complète pour obtenir le problème modèle. Les mélieux poreux sont décrits pour introduire les notions de base et la phénoménologie des écoulements poreux. Les méthodes de changement d'échelle sont discutées en tant qu'outils de modélisation pour obtenir les relations fondamentales des écoulements dans les milieux poreux. En particulier, elles peuvent servir à mieux comprendre l'utilisation de l'équation de Richards et constituent des perspectives de travail intéressantes. La dérivation de l'équation de Richards est donnée pour esquisser les limites du modèle. Des sections sont consacrées aux propriétés hydrauliques et à la condition aux limites de suintement. Ce sont des approximationx de processus spécifiques aux milieux poreux. En outre, elles contiennent des non-linéarités responsables des difficultés numériques de l'équation de Richards.

Dans le Chapitre II, l'équation de Richards est discrétisée. Après avoir introduit les avantages des méthodes GD, la discrétisation en espace est détaillée. L'utilisation d'une pénalisation faible pour imposer les conditions aux limites est mise en évidence. La discrétisation en temps est effectuée par des méthodes implicites appelées formule de différentiation rétrograde (FDR). La propriété de conservation de la masse des méthodes de GD est présentée. Le problème discret non-linéaire est ensuite résolu par un processus itératif basé sur la méthode de Newton-Raphson ou la méthode du point fixe. Un pas de temps adaptatif heuristique est utilisé pour rendre la convergence non-linéaire plus robuste. À la fin de ce chapitre, des choix de mise en œuvre sont fournis étant donné que les méthodes GD offrent plusieurs possibilités.

Le Chapitre III est consacré à la stratégie adaptative proposée pour améliorer la résolution de l'équation de Richards. Le RMA par blocs est présenté et justifié pour l'équation de Richards avec une utilisation de maillages généraux. Ensuite, un indicateur d'erreur basé sur le gradient et un indicateur d'erreur basé sur l'estimation sont introduits et motivés sur leur habileté à décider de l'adaptation du maillage. Ils sont conçus pour capturer les fronts raides. Quelques variations autour du schéma numérique sont également proposées. En particulier, les sauts et les flux sont traités par la méthode GDP pour mieux approcher les fronts raides grâce à une discontinuité.

Plusieurs tests numériques sont effectués dans le Chapitre IV. Les expériences numériques permettent de se faire une idée des méthodes GD. Le modèle est validé par rapport à des cas-tests d'infiltration verticale 1D impliquant différentes propriétés hydrauliques et conditions aux limites. Certains comportements spécifiques de l'équation de Richards sont notés et les performances de l'adaptation du maillage et du cadre pondéré pour capturer les fronts de saturation sont démontrées. Le cas de Tracy fournit une solution exacte qui est utilisée pour évaluer la convergence et l'efficacité des indicateurs d'erreur. Le code est testé sur des expériences de laboratoire impliquant du suintement, une recharge et des drai-

nages de nappe phréatique. Enfin, la stratégie de modélisation est testée sur un cas difficile à grande échelle avec un ensemble de données expérimentales : la saturation d'un barrage multi-matériaux.

Dans le Chapitre V, des applications sont menées pour les écoulements souterrains des plages sableuses. Après avoir présenté les notions de système de plage impliquant le forçage des vagues et la circulation des écoulements souterrains, des comparaisons sont effectuées pour trois cas. Une plage idéalisée est considérée pour voir si le modèle est capable de rendre les principales caractéristiques des écoulements souterrains de plage observées dans les expériences disponibles dans la littérature.

Synthèse conclusive

Dans cette thèse, le but principal consistait à développer des simulations précises, robustes et efficaces de l'équation de Richards pour modéliser les écoulements souterrains des plages. À cette fin, une nouvelle stratégie a été formulée pour la solution de l'équation de Richards sur la base de la méthode GD avec des FDR et un RMA.

L'approximation de l'équation de Richards peut atteindre un ordre élevé à la fois en espace et en temps, traiter des conditions de suintement, des conditions aux limites dynamiques, des milieux poreux hétérogènes et des fronts de saturation/désaturation raides. Les questions exposées précédemment ont été abordées et discutées plus en détail tout au long de la thèse, de sorte que les remarques suivantes peuvent être tirées en guise de conclusion :

- L'équation de Richards a été dérivée en soulignant des hypothèses claires pour la modélisation ;
- La condition aux limites de suintement a été directement incorporée dans la formulation GD faible, ce qui rend son traitement naturel et simple.
- Le pas de temps adaptatif a permis au solveur itératif non-linéaire de converger, ce qui lui confère une grande robustesse ;
- L'adaptation du maillage a été utilisée pour contrôler les erreurs spatiales de l'équation de Richards en capturant les fronts de saturation dynamiques grâce à un indicateur d'erreur *a posteriori* basé sur l'estimation ;
- Le RAM a été basé sur une structure par blocs pour délimiter de grandes régions où les fronts de saturation se déplacent, afin de gagner du temps de calcul en évitant un raffinement systématique. Elle prépare également une parallélisation future ;
- Les oscillations des fronts de saturation raides de l'équation de Richards ont été interprétées dans le contexte d'une équation de convection-diffusion. Les fronts de saturation ont été considérés comme des couches internes en raison de la diffusivité variable non-linéaire, voire dégénérée, ainsi que de l'hétérogénéité et de l'anisotropie des matériaux. Un nombre de Péclet local a été présenté pour mettre en évidence les régions problématiques. Une méthode GDP a été utilisée pour permettre des sauts dans la solution qui approximent mieux les couches internes raides. L'adaptation du maillage et le cadre pondéré fonctionnent en synergie pour capturer et résoudre les fronts de saturation raides grâce à l'indicateur d'erreur *a posteriori* proposé.

Le Chapitre I a apporté des éclaircissements sur les possibilités de modélisation de l'équations de Richards. Cette équation parabolique non-linéaire simule un système air-eau, mais en ne considérant que la phase liquide. Ainsi, elle est très utile pour modéliser les écoulements en milieux poreux variablement saturés, mais il existe des limitations comme la phase gazeuse qui doit rester continuellement connectée. Les propriétés hydrauliques supposent que les phases sont à l'équilibre et ne tiennent pas

compte de l'hystérésis. La modélisation de l'équation de Richards reste riche en prenent en compte les phénomènes de gravité et les effets capillaires. Ce double comportement associé à la non-linéarité rend l'équation de Richards difficile à résoudre numériquement pour un large éventail d'applications.

Le Chapitre II a présenté les méthodes numériques utilisées pour l'équation de Richards. La discrétisation de l'espace a été réalisée par des méthodes de GD. La formulation faible de la GD donne la flexibilité nécessaire pour traiter des maillages hybrides non-conformes non-structurés, des ordres d'approximation élevés et des conditions aux limites spéciales. Des FDR sont utilisées pour la discrétisation en temps afin d'atteindre un ordre élevé et d'assurer la stabilité. La résolution non-linéaire repose sur la méthode du point fixe dont la convergence est contrôlée par un pas de temps adaptatif heuristique. Certains choix d'implémentation dans le code *Rivage* ont été présentés pour les solveurs, les fonctions de base ou l'intégration numérique.

Le Chapitre III a proposé la stratégie adaptative basée sur le RAM. Une structure basée sur des blocs a permis d'économiser du temps de calcul et de préparer une future parallélisation. L'adaptation du maillage est déterminée par des indicateurs d'erreur appuyés soit par la physique du problème, soit par une estimation *a posteriori*. Ils sont combinés avec le cadre GDP pour contrôler les sauts et les flux dans la solution. Les fronts de saturation peuvent être si raides qu'une discontinuité approxime mieux la solution de l'équation de Richards.

Le Chapitre IV a utilisé plusieurs cas-tests. Diverses infiltrations 1D dans une colonne de sol ont permis de saisir la dynamique de l'équation de Richards, de valider le code et de tester les techniques de capture par l'adaptation du maillage et le cadre pondéré. En particulier, des sauts ont été introduits par la pénalité pondérée qui empêche la propagation des oscillations lorsque le front est raide. L'indicateur d'erreur basé sur l'estimation s'est révélé très performant pour conduire l'adaptation du maillage efficacement. En outre, un schéma temporel d'ordre élevé a été nécessaire pour atteindre une précision élevée lorsque les erreurs spatiales sont faibles. Un pas de temps adaptatif a été nécessaire pour rendre la convergence non-linéaire fiable. Le code a permis de simuler diverses dynamiques (recharge/drainage) d'expériences de laboratoire. La saturation du barrage multi-matériaux de La Verne est un cas exigeant présentant des données expérimentales. Une comparaison a montré un très bon accord du modèle avec la dynamique du système. Qui plus est, elle a démontré la capacité de la stratégie à traiter des fronts raides causés par l'hétérogénéité ou la diffusivité variable non-linéaire. L'adaptation du maillage a permis de capturer des fronts avec l'indicateur d'erreur basé sur l'estimation et la taille des sauts ajustée par la méthode GDP. Sans apporter de solution parfaite, cette stratégie est très rentable par rapport au raffinement du maillage requis pour améliorer les résultats.

Le Chapitre V a appliqué la méthode aux écoulements souterrains dans les plages sableuses. Le modèle basé sur l'équation de Richards a pu rendre les principales caractéristiques observées expérimentalement même si seul un couplage forcé est utilisé. Ceci est démontré par un cas idéalisé de plage.

Cette stratégie conduit à une solution conservative, efficace et fiable de l'équation de Richards. Cependant, la méthode contient des paramètres heuristiques de sorte que des questions restent ouvertes concernant l'analyse numérique, la prise de décision *hp*, la convergence non-linéaire ou le bilan d'erreurs. Tous ces aspects ont un grand potentiel d'amélioration et seront étudiés dans de futurs travaux afin d'optimiser la résolution de l'équation de Richards.

Les suggestions suivantes sont fournies pour les recherches futures :

- Un travail de validation et de confrontation est en cours avec des cas impliquant des expériences en laboratoire et à grande échelle pour les écoulements souterrains dans les plages sableuses. Une application à une expérience de terrain sur la plage de Rousty est visée pour apporter de nouvelles connaissances sur les écoulements souterrains de jet de rive ;

- Il serait intéressant de poursuivre la méthode de la prise de moyenne volumique afin de dégager des hypothèses claires sur le régime d'écoulement et le milieu poreux. Tout d'abord, cela contribuera au domaine des méthode de changment d'échelle en proposant une analyse dimensionnelle basée sur différentes longueurs caractéristiques. Deuxièmement, cela précisera la plage de validité de la relation fondamentale modélisant l'écoulement du milieu poreux, qui est ensuite utilisé dans l'équation de Richards. Enfin, la prise de moyenne volumique est également un outil puissant pour élaborer des modèles de couplage entre les écoulements souterrains et de surface. Cela contribuera à construire un couplage plus élaboré qu'un forçage ;
- Les paramètres de pénalisation assurent la stabilité des méthodes GD et imposent les conditions aux limites. Cependant, ils peuvent également dégrader les propriétés numériques avec des valeurs arbitrairement grandes, en particulier le conditionnement, ce qui entraîne des instabilités. De plus, ils offrent un moyen d'incorporer dans le schéma une connaissance partielle de la solution. Une analyse numérique et d'autres tests seraient effectués afin de dériver une définition adaptée au problème ou, du moins, des bornes appropriées ;
- Le pas de temps adaptatif contrôle la convergence itérative non-linéaire. Cependant, la méthode n'empêche pas le pas de temps de diminuer au point de rendre le calcul impossible. Qui plus est, la méthode est heuristique et il est difficile de savoir si du temps de calcul est économisé puisque le facteur d'amplication/réduction est un paramètre défini par l'utilisateur. Une possibilité serait d'utiliser un estimateur d'erreur temporelle. Une autre possibilité serait d'étudier la convergence de la procédure non-linéaire qui dépend du pas de temps et d'en déduire une méthode. Un pas de temps local pourrait être une idée pour économiser du temps de calcul mais la mise en œuvre est difficile avec des schémas implicites ;
- La sélection de l'ordre temporel est pour l'instant fixée dans le code. Cette partie devrait être entièrement automatique et fournir la meilleure FDR pour garantir la stabilité. Cela se fait généralement au moyen d'estimateurs d'erreurs ou d'une analyse de stabilité ;
- L'adaptation *hp* serait une solution efficace pour la solution de l'équation de Richards en effectuant un raffinement du maillage ou un enrichissement de l'ordre suivant la meilleure possibilité. De nombreuses techniques sont déjà présentes dans la littérature pour prendre des décisions de type *hp*. Certaines remarques dans cette thèse laissent penser que les sous-parties de l'indicateur d'erreur basé sur l'estimation peuvent conduire à la décision entre l'adaptation *h* et l'adaptation *p*. Même si cette adaptation *hp* ne sera pas optimale, la technique serait très simple et rentable. D'autres tests sont nécessaires et une véritable analyse d'estimation d'erreur *a posteriori* pourrait être utile ;
- L'adaptabilité est un problème pour la fréquence de raffinement et la sélection des éléments à raffiner. Un seuillage automatique basé sur une fonction de distribution serait une possibilité pour décider des éléments à raffiner/déraffiner. La dérivation de la vitesse des fronts de saturation serait un moyen de définir la fréquence de raffinement en fonction de la taille du bloc ;
- La parallélisation et la 3D sont des extensions naturelles du code *Rivage* pour traiter des cas plus complexes et gagner du temps de calcul. Même si le code est préparé pour cela (structure basée sur des blocs conduisant à la décomposition de domaine et implémentation générale), ce ne sera pas simple car la 3D implique de nombreuses géométries d'éléments et que le schéma temporel est implicite. En outre, l'ajout d'un module pour traiter des équations hyperboliques serait une possibilité à étudier afin d'incorporer des modèles d'écoulement de surface et effectuer un couplage adéquat.

Plusieurs parties de la thèse ont été présentées lors de conférences et de séminaires dont certains ont

donné lieu à des actes. Un article a été proposé pour publication à un journal dans lequel la stratégie adaptative GD est décrite afin de simuler l'équation de Richards. Un autre article est prévu pour présenter et discuter des applications des écoulements souterrains aux plages sableuses. Ces travaux sont répertoriés dans la List of Publications, page 145.

Lists

List of Figures

List of Tables

List of Publications

Journal articles:

- Jean-Baptiste Clément, Damien Sous, Frédéric Golay, Mehmet Ersoy,
  Wave-driven Groundwater Flows in Sandy Beaches: A Richards Equation-based Model,
  Journal of Coastal Research, 95 (sp1), 1047-1051, 26 May 2020.
  DOI: 10.2112/SI95-204.1. [256]

Conference proceedings:

- Jean-Baptiste Clément, Frédéric Golay, Mehmet Ersoy, Damien Sous,
  Adaptive Discontinuous Galerkin Method for Richards Equation,
  Topical Problems of Fluid Mechanics 2020, Prague, Czech Republic, 27-34, February 2020.
  DOI: 10.14311/TPFM.2020.004. [257]

- Jean-Baptiste Clément, Mehmet Ersoy, Frédéric Golay, Damien Sous,
  Discontinuous Galerkin Method for Steady-State Richards Equation,
  Topical Problems of Fluid Mechanics 2019, Prague, Czech Republic, 53-62, February 2019.
  DOI: 10.14311/TPFM.2019.008. [258]

Submitted or working papers:

- Jean-Baptiste Clément, Frédéric Golay, Mehmet Ersoy, Damien Sous,
  An adaptive strategy for discontinuous Galerkin simulations of Richards' equation.
  Submission in progress.

- Jean-Baptiste Clément, Damien Sous, Frédéric Golay, Mehmet Ersoy,
  A Richards' equation based model for wave-resolving variable-saturation groundwater beach dynamics.
  Writing in progress.

Talks:

- International Coastal Symposium 2020, *Wave-driven Groundwater Flows in Sandy Beaches: A Richards Equation-based Model*, Seville, Spain. Re-scheduled to May 2021 due to COVID-19 pandemic.

- WCCM-ECCOMAS Congress 2020, hp-*Adaptation for Discontinuous Galerkin Methods Applied to Variably-Saturated Flows*, Paris, France. Re-scheduled to January 2021 due to COVID-19 pandemic.

- TPFM 2020, *Adaptive Discontinuous Galerkin Method for Richards Equation*, Prague, Czech Republic, 20 February 2020.

- Workshop GLADYS New Eve 2020, *Simulation numérique des écoulements insaturés dans les plages sableuses*, Université de Montpellier/MUSE, France, 7 January 2020.

- Current Problems in Numerical Analysis, seminar from Institute of Mathematics of the Czech Academy of Sciences, *Richards equation for flows in sandy beaches: from modelling to Discontinuous Galerkin simulation*, Prague, Czech Republic, 25 october 2019.

- TPFM 2019, *Discontinuous Galerkin Method for Steady-State Richards Equation*, Prague, Czech Republic, 20 February 2019.

- IMATH laboratory seminar, *Flows in unsaturated porous media by discontinuous Galerkin method*, La Garde, France, 7 February 2019.

Bibliography

[1] A. Luijendijk, G. Hagenaars, R. Ranasinghe, F. Baart, G. Donchyts, S. Aarninkhof, The State of the World's Beaches, Scientific Reports 8 (1). doi:10.1038/s41598-018-24630-6.

[2] M. I. Vousdoukas, R. Ranasinghe, L. Mentaschi, T. A. Plomaritis, P. Athanasiou, A. Luijendijk, L. Feyen, Sandy coastlines under threat of erosion, Nature Climate Change 10 (3) (2020) 260–263. doi:10.1038/s41558-020-0697-0.

[3] K. Steenhauer, D. Pokrajac, T. O'Donoghue, G. A. Kikkert, Subsurface processes generated by bore-driven swash on coarse-grained beaches, Journal of Geophysical Research 116 (C4). doi:10.1029/2010jc006789.

[4] I. L. Turner, G. C. Rau, M. J. Austin, M. S. Andersen, Groundwater fluxes and flow paths within coastal barriers: Observations from a large-scale laboratory experiment (BARDEX II), Coastal Engineering 113 (2016) 104–116. doi:10.1016/j.coastaleng.2015.08.004.

[5] J. W. Heiss, J. A. Puleo, W. J. Ullman, H. A. Michael, Coupled surface-subsurface hydrologic measurements reveal infiltration, recharge, and discharge dynamics across the swash zone of a sandy beach, Water Resources Research 51 (11) (2015) 8834–8853. doi:10.1002/2015wr017395.

[6] D. Sous, L. Petitjean, F. Bouchette, V. Rey, S. Meulé, F. Sabatier, K. Martins, Field evidence of swash groundwater circulation in the microtidal rousty beach, France, Advances in Water Resources 97 (2016) 144–155. doi:10.1016/j.advwatres.2016.09.009.

[7] D. Sous, L. Petitjean, F. Bouchette, V. Rey, S. Meulé, F. Sabatier, Groundwater fluxes within sandy beaches: swash-driven flow vs cross-barrier gradients, in: Coastal Dynamics 2017, 2017.

[8] L. Li, D. A. Barry, Wave-induced beach groundwater flow, Advances in Water Resources 23 (4) (2000) 325–337. doi:10.1016/s0309-1708(99)00032-9.

[9] L. Li, D. A. Barry, C. B. Pattiaratchi, G. Masselink, BeachWin: modelling groundwater effects on swash sediment transport and beach profile changes, Environmental Modelling & Software 17 (3) (2002) 313–320. doi:10.1016/s1364-8152(01)00066-4.

[10] H. Darcy, Les fontaines publiques de la ville de Dijon, Librairie des corps impériaux des ponts et des chaussées et des mines, 1856.

[11] E. Buckingham, Studies on the movement of soil moisture, bulletin Edition, Vol. 38, Washington, Govt. Print. Off., 1907.

[12] L. F. Richardson, Weather Prediction by Numerical Process, Cambridge University Press, 1922.

[13] L. A. Richards, Capillary conduction of liquids through porous mediums, Physics 1 (5) (1931) 318–333. doi:10.1063/1.1745010.

[14] J. Rubin, Theoretical Analysis of Two-Dimensional, Transient Flow of Water in Unsaturated and Partly Unsaturated Soils, Soil Science Society of America Journal 32 (5) (1968) 607–615. doi:10.2136/sssaj1968.03615995003200050013x.

[15] R. L. Cooley, A Finite Difference Method for Unsteady Flow in Variably Saturated Porous Media: Application to a Single Pumping Well, Water Resources Research 7 (6) (1971) 1607–1625. doi:10.1029/wr007i006p01607.

[16] H. Wilhelm Alt, S. Luckhaus, Quasilinear elliptic-parabolic differential equations, Mathematische Zeitschrift 183 (3) (1983) 311–341. doi:10.1007/BF01176474.

[17] C. Paniconi, M. Putti, A comparison of Picard and Newton iteration in the numerical solution of multidimensional variably saturated flow problems, Water Resources Research 30 (12) (1994) 3357–3374. doi:10.1029/94wr02046.

[18] F. Lehmann, P. Ackerer, Comparison of Iterative Methods for Improved Solutions of the Fluid Flow Equation in Partially Saturated Porous Media, Transport in Porous Media 31 (3) (1998) 275–292. doi:10.1023/a:1006555107450.

[19] R. B. Thoms, R. L. Johnson, R. W. Healy, User's guide to the Variably Saturated Flow (VSF) process to MODFLOW (2006). doi:10.3133/tm6a18.

[20] C. T. Miller, C. Abhishek, M. W. Farthing, A spatially and temporally adaptive solution of Richards' equation, Advances in Water Resources 29 (4) (2006) 525–545. doi:10.1016/j.advwatres.2005.06.008.

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

[22] A. Szymkiewicz, Modelling Water Flow in Unsaturated Porous Media, Springer Berlin Heidelberg, 2013. doi:10.1007/978-3-642-23559-7.

[23] M. W. Farthing, F. L. Ogden, Numerical Solution of Richards' Equation: A Review of Advances and Challenges, Soil Science Society of America Journal 81 (6) (2017) 1257–1269. doi:10.2136/sssaj2017.02.0058.

[24] Y. Zha, J. Yang, J. Zeng, C.-H. M. Tso, W. Zeng, L. Shi, Review of numerical solution of Richardson–Richards equation for variably saturated flow in soils, Wiley Interdisciplinary Reviews: Water 6 (5). doi:10.1002/wat2.1364.

[25] J. A. G. Cooper, G. Masselink, G. Coco, A. D. Short, B. Castelle, K. Rogers, E. Anthony, A. N. Green, J. T. Kelley, O. H. Pilkey, D. W. T. Jackson, Sandy beaches can survive sea-level rise, Nature Climate Change 10 (11) (2020) 993–995. doi:10.1038/s41558-020-00934-2.

[26] B. Rivière, Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations, Society for Industrial and Applied Mathematics, 2008. doi:10.1137/1.9780898717440.

[27] V. Dolejší, M. Feistauer, Discontinuous Galerkin Method, Springer International Publishing, 2015. doi:10.1007/978-3-319-19267-3.

[28] W. Klieber, B. Rivière, Adaptive simulations of two-phase flow by discontinuous Galerkin methods, Computer Methods in Applied Mechanics and Engineering 196 (1-3) (2006) 404–419. doi:10.1016/j.cma.2006.05.007.

[29] P. Bastian, A fully-coupled discontinuous Galerkin method for two-phase flow in porous media with discontinuous capillary pressure, Computational Geosciences 18 (5) (2014) 779–796. doi:10.1007/s10596-014-9426-y.

[30] F. List, F. A. Radu, A study on iterative methods for solving Richards' equation, Computational Geosciences 20 (2) (2016) 341–353. doi:10.1007/s10596-016-9566-3.

[31] H. Li, M. W. Farthing, C. N. Dawson, C. T. Miller, Local discontinuous Galerkin approximations to Richards' equation, Advances in Water Resources 30 (3) (2007) 555–575. doi:10.1016/j.advwatres.2006.04.011.

[32] A. Ern, A. F. Stephansen, P. Zunino, A discontinuous Galerkin method with weighted averages for advection-diffusion equations with locally small and anisotropic diffusivity, IMA Journal of Numerical Analysis 29 (2) (2008) 235–256. doi:10.1093/imanum/drm050.

[33] D. A. Di Pietro, A. Ern, J.-L. Guermond, Discontinuous Galerkin Methods for Anisotropic Semidefinite Diffusion with Advection, SIAM Journal on Numerical Analysis 46 (2) (2008) 805–831. doi:10.1137/060676106.

[34] J. Proft, B. Rivière, Analytical and numerical study of diffusive fluxes for transportequations with near-degenerate coefficients, Tech. Rep. TR-MATH 06-07, Universityof Pittsburgh (2006).

[35] J. Proft, B. Rivière, Discontinuous Galerkin methods for convectiondiffusion equations with varying and vanishing diffusivity, International Journal ofNumerical Analysis and Modeling 6 (4) (2009) 533–561.

[36] J. Bear, Dynamics of fluids in porous media, American Elsevier Publishing Company, 1972.

[37] R. Helmig, Multiphase Flow and Transport Processes in the Subsurface, Springer-Verlag Berlin Heidelberg, 1997.

[38] U. Hornung (Ed.), Homogenization and Porous Media, Springer New York, 1997. doi:10.1007/978-1-4612-1920-0.

[39] G. Allaire, Homogenization and Two-Scale Convergence, SIAM Journal on Mathematical Analysis 23 (6) (1992) 1482–1518. doi:10.1137/0523084.

[40] G. Allaire, Shape Optimization by the Homogenization Method, Springer New York, 2002. doi:10.1007/978-1-4684-9286-6.

[41] J. L. Auriault, Heterogeneous medium. is an equivalent macroscopic description possible?, International Journal of Engineering Science 29 (7) (1991) 785–795. doi:10.1016/0020-7225(91)90001-j.

[42] J. L. Auriault, P. M. Adler, Taylor dispersion in porous media: Analysis by multiple scale expansions, Advances in Water Resources 18 (4) (1995) 217–226. doi:10.1016/0309-1708(95)00011-7.

[43] S. Whitaker, Flow in porous media I: A theoretical derivation of Darcy's law, Transport in Porous Media 1 (1) (1986) 3–25. doi:10.1007/bf01036523.

[44] S. Whitaker, Flow in porous media II: The governing equations for immiscible, two-phase flow, Transport in Porous Media 1 (2) (1986) 105–125. doi:10.1007/bf00714688.

[45] S. Whitaker, Flow in porous media III: Deformable media, Transport in Porous Media 1 (2) (1986) 127–154. doi:10.1007/bf00714689.

[46] S. Whitaker, The Method of Volume Averaging, Springer Netherlands, 1999. doi:10.1007/978-94-017-3389-2.

[47] M. Quintard, S. Whitaker, Transport in ordered and disordered porous media I: The cellular average and the use of weighting functions, Transport in Porous Media 14 (2) (1994) 163–177. doi:10.1007/bf00615199.

[48] M. Quintard, S. Whitaker, Transport in ordered and disordered porous media II: Generalized volume averaging, Transport in Porous Media 14 (2) (1994) 179–206. doi:10.1007/bf00615200.

[49] M. Quintard, S. Whitaker, Transport in ordered and disordered porous media III: Closure and comparison between theory and experiment, Transport in Porous Media 15 (1) (1994) 31–49. doi:10.1007/bf01046157.

[50] M. Quintard, S. Whitaker, Transport in ordered and disordered porous media IV: Computer generated porous media for three-dimensional systems, Transport in Porous Media 15 (1) (1994) 51–70. doi:10.1007/bf01046158.

[51] M. Quintard, S. Whitaker, Transport in ordered and disordered porous media v: Geometrical results for two-dimensional systems, Transport in Porous Media 15 (2) (1994) 183–196. doi:10.1007/bf00625516.

[52] W. G. Gray, S. M. Hassanizadeh, Paradoxes and Realities in Unsaturated Flow Theory, Water Resources Research 27 (8) (1991) 1847–1854. doi:10.1029/91wr01259.

[53] S. P. Neuman, Theoretical derivation of Darcy's law, Acta Mechanica 25 (3-4) (1977) 153–170. doi:10.1007/bf01376989.

[54] C. M. Marle, Écoulements monophasiques en milieu poreux, Revue de l'Institut Français du Pétrole 22 (10) (1967) 1471–1509.

[55] C. M. Marle, Les écoulements polyphasiques en milieu poreux: de l'échelle des pores à l'échelle macroscopique, Annales des Mines (5-6) (1984) 51–56.

[56] G. Dagan, The generalization of Darcy's Law for nonuniform flows, Water Resources Research 15 (1) (1979) 1–7. doi:10.1029/wr015i001p00001.

[57] D. Zhang, Stochastic Methods for Flow in Porous Media, Elsevier, 2002. doi:10.1016/b978-0-12-779621-5.x5000-1.

[58] G. Matheron, Genèse et signification énergétique de la loi de Darcy, Revue de l'Institut Français du Pétrole 21 (11) (1966) 1697–1706.

[59] G. Matheron, L'émergence de la loi de Darcy, Annales des Mines (5-6) (1984) 11–16.

[60] Y. Davit, C. G. Bell, H. M. Byrne, L. A. C. Chapman, L. S. Kimpton, G. E. Lang, K. H. L. Leonard, J. M. Oliver, N. C. Pearson, R. J. Shipley, S. L. Waters, J. P. Whiteley, B. D. Wood, M. Quintard, Homogenization via formal multiscale asymptotics and volume averaging: How do the two techniques compare?, Advances in Water Resources 62 (2013) 178–206. doi:10.1016/j.advwatres.2013.09.006.

[61] G. S. Beavers, D. D. Joseph, Boundary conditions at a naturally permeable wall, Journal of Fluid Mechanics 30 (1) (1967) 197–207. doi:10.1017/s0022112067001375.

[62] P. G. Saffman, On the Boundary Condition at the Surface of a Porous Medium, Studies in Applied Mathematics 50 (2) (1971) 93–101. doi:10.1002/sapm197150293.

[63] K. Fujisawa, A. Murakami, Numerical analysis of coupled flows in porous and fluid domains by the Darcy-Brinkman equations, Soils and Foundations 58 (5) (2018) 1240–1259. doi:10.1016/j.sandf.2018.07.003.

[64] H.-H. Liu, Generalization of the Darcy-Buckingham Law: Optimality and Water Flow in Unsaturated Media, in: Fluid Flow in the Subsurface, Springer International Publishing, 2016, pp. 45–102. doi:10.1007/978-3-319-43449-0_2.

[65] I. J. Losada, J. L. Lara, M. del Jesus, Modeling the Interaction of Water Waves with Porous Coastal Structures, Journal of Waterway, Port, Coastal, and Ocean Engineering 142 (6) (2016) 03116003. doi:10.1061/(asce)ww.1943-5460.0000361.

[66] F. A. Howes, S. Whitaker, The spatial averaging theorem revisited, Chemical Engineering Science 40 (8) (1985) 1387–1392. doi:10.1016/0009-2509(85)80078-6.

[67] V. Baron, Y. Coudière, P. Sochala, Adaptive multistep time discretization and linearization based on a posteriori error estimates for the Richards equation, Applied Numerical Mathematics 112 (2017) 104–125. doi:10.1016/j.apnum.2016.10.005.

[68] S. Irmay, On the hydraulic conductivity of unsaturated soils, Transactions, American Geophysical Union 35 (3) (1954) 463. doi:10.1029/tr035i003p00463.

[69] W. R. Gardner, Some Steady-State Solutions of the Unsaturated Moisture Flow Equation with Application to Evaporation from a Water Table, Soil Science 85 (4) (1958) 228–232. doi:10.1097/00010694-195804000-00006.

[70] G. Vachaud, J.-L. Thony, Hysteresis During Infiltration and Redistribution in a Soil Column at Different Initial Water Contents, Water Resources Research 7 (1) (1971) 111–127. doi:10.1029/wr007i001p00111.

[71] Y. Mualem, A new model for predicting the hydraulic conductivity of unsaturated porous media, Water Resources Research 12 (3) (1976) 513–522. doi:10.1029/wr012i003p00513.

[72] M. T. van Genuchten, A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, Soil Science Society of America Journal 44 (5) (1980) 892–898. doi:10.2136/sssaj1980.03615995004400050002x.

[73] N. T. Burdine, Relative Permeability Calculations From Pore Size Distribution Data, Journal of Petroleum Technology 5 (03) (1953) 71–78. doi:10.2118/225-g.

[74] T. Vogel, M. T. van Genuchten, M. Cislerova, Effect of the shape of the soil hydraulic functions near saturation on variably-saturated flow predictions, Advances in Water Resources 24 (2) (2000) 133–144. doi:10.1016/s0309-1708(00)00037-3.

[75] B. Schweizer, The Richards equation with hysteresis and degenerate capillary pressure, Journal of Differential Equations 252 (10) (2012) 5594–5612. doi:10.1016/j.jde.2012.01.026.

[76] A. S. Krämer, A discontinuous Galerkin scheme for the Richards equation: on dynamic capillarity effects in porous media, Master's thesis, Eindhoven University of Technology (2016).

[77] C. Scudeler, C. Paniconi, D. Pasetto, M. Putti, Examination of the seepage face boundary condition in subsurface and coupled surface/subsurface hydrological models, Water Resources Research 53 (3) (2017) 1799–1819. doi:10.1002/2016wr019277.

[78] P. Sochala, A. Ern, S. Piperno, Mass conservative BDF-discontinuous Galerkin/explicit finite volume schemes for coupling subsurface and overland flows, Computer Methods in Applied Mechanics and Engineering 198 (27-29) (2009) 2122–2136. doi:10.1016/j.cma.2009.02.024.

[79] P. A. Forsyth, M. C. Kropinski, Monotonicity Considerations for Saturated–Unsaturated Subsurface Flow, SIAM Journal on Scientific Computing 18 (5) (1997) 1328–1354. doi:10.1137/s1064827594265824.

[80] C. E. Kees, M. W. Farthing, C. N. Dawson, Locally conservative, stabilized finite element methods for variably saturated flow, Computer Methods in Applied Mechanics and Engineering 197 (51-52) (2008) 4610–4625. doi:10.1016/j.cma.2008.06.005.

[81] S. Bonelli, F. Golay, O. Débordes, Résolution auto-adaptative par éléments finis de problèmes de diffusion fortement non linéaires, Revue Européenne des Éléments Finis 2 (4) (1993) 495–515. doi:10.1080/12506559.1993.10511094.

[82] B. H. Gilding, Qualitative mathematical analysis of the Richards equation, Transport in Porous Media 6 (5-6). doi:10.1007/bf00137854.

[83] J.-G. Caputo, Y. A. Stepanyants, Front Solutions of Richards' Equation, Transport in Porous Media 74 (1) (2007) 1–20. doi:10.1007/s11242-007-9180-x.

[84] P. Sochala, Méthodes numériques pour les écoulements souterrains et couplage avec le ruissellement, Ph.D. thesis, École Nationale des Ponts et Chaussées, thèse de doctorat dirigée par Ern, Alexandre et Piperno, Serge Mathématiques. Informatique Paris Est 2008 (2008).
URL http://www.theses.fr/2008ENPC0813

[85] M. A. Celia, E. T. Bouloutas, R. L. Zarba, A general mass-conservative numerical solution for the unsaturated flow equation, Water Resources Research 26 (7) (1990) 1483–1496. doi:10.1029/wr026i007p01483.

[86] L. Pan, A. W. Warrick, P. J. Wierenga, Finite element methods for modeling water flow in variably saturated porous media: Numerical oscillation and mass-distributed schemes, Water Resources Research 32 (6) (1996) 1883–1889. doi:10.1029/96wr00753.

[87] S.-H. Ju, K.-J. S. Kung, Mass types, element orders and solution schemes for the richards equation, Computers & Geosciences 23 (2) (1997) 175–187. doi:10.1016/s0098-3004(97)85440-4.

[88] A. Younes, M. Fahs, B. Belfort, Monotonicity of the cell-centred triangular MPFA method for saturated and unsaturated flow in heterogeneous porous media, Journal of Hydrology 504 (2013) 132–141. doi:10.1016/j.jhydrol.2013.09.041.

[89] B. Belfort, A. Younes, M. Fahs, F. Lehmann, On equivalent hydraulic conductivity for oscillation–free solutions of Richard's equation, Journal of Hydrology 505 (2013) 202–217. doi:10.1016/j.jhydrol.2013.09.047.

[90] P. M. Gresho, R. L. Lee, Don't suppress the wiggles—They're telling you something!, Computers & Fluids 9 (2) (1981) 223–253. doi:10.1016/0045-7930(81)90026-8.

[91] P. A. Forsyth, Y. S. Wu, K. Pruess, Robust numerical methods for saturated-unsaturated flow with dry initial conditions in heterogeneous media, Advances in Water Resources 18 (1) (1995) 25–38. doi:10.1016/0309-1708(95)00020-j.

[92] C. S. Woodward, C. N. Dawson, Analysis of Expanded Mixed Finite Element Methods for a Nonlinear Parabolic Equation Modeling Flow into Variably Saturated Porous Media, SIAM Journal on Numerical Analysis 37 (3) (2000) 701–724. doi:10.1137/s0036142996311040.

[93] A. Dogan, L. H. Motz, Saturated-Unsaturated 3D Groundwater Model. II: Verification and Application, Journal of Hydrologic Engineering 10 (6) (2005) 505–515. doi:10.1061/(asce)1084-0699(2005)10:6(505).

[94] G. Manzini, S. Ferraris, Mass-conservative finite volume methods on 2-D unstructured grids for the Richards' equation, Advances in Water Resources 27 (12) (2004) 1199–1215. doi:10.1016/j.advwatres.2004.08.008.

[95] L. Orgogozo, N. Renon, C. Soulaine, F. Hénon, S. K. Tomer, D. Labat, O. S. Pokrovsky, M. Sekhar, R. Ababou, M. Quintard, An open source massively parallel solver for Richards equation: Mechanistic modelling of water fluxes at the watershed scale, Computer Physics Communications 185 (12) (2014) 3358–3371. doi:10.1016/j.cpc.2014.08.004.

[96] D. Caviedes-Voullième, J. Murillo, P. Garćia-Navarro, Numerical simulation of groundwater-surface interactions by external coupling of the 3D Richards equation and the full 2D shallow-water equations, in: XIX International Conference on Water Resources, 2012.

[97] J. Šimůnek, M. T. van Genuchten, M. Šejna, Development and Applications of the HYDRUS and STANMOD Software Packages and Related Codes, Vadose Zone Journal 7 (2) (2008) 587–600. doi:10.2136/vzj2007.0077.

[98] D. Kavetski, P. Binning, S. W. Sloan, Adaptive backward Euler time stepping with truncation error control for numerical modelling of unsaturated fluid flow, International Journal for Numerical Methods in Engineering 53 (6) (2001) 1301–1322. doi:10.1002/nme.329.

[99] H.-J. G. Diersch, P. Perrochet, On the primary variable switching technique for simulating unsaturated–saturated flows, Advances in Water Resources 23 (3) (1999) 271–301. doi:10.1016/s0309-1708(98)00057-8.

[100] P. Knabner, E. Schneid, Adaptive Hybrid Mixed Finite Element Discretization of Instationary Variably Saturated Flow in Porous Media, in: Lecture Notes in Computational Science and Engineering, Springer Berlin Heidelberg, 2002, pp. 37–44. doi:10.1007/978-3-642-55919-8_4.

[101] G. Chavent, J. E. Roberts, A unified physical presentation of mixed, mixed-hybrid finite elements and standard finite difference approximations for the determination of velocities in waterflow problems, Advances in Water Resources 14 (6) (1991) 329–348. doi:10.1016/0309-1708(91)90020-o.

[102] M. Fahs, A. Younes, F. Lehmann, An easy and efficient combination of the Mixed Finite Element Method and the Method of Lines for the resolution of Richards' Equation, Environmental Modelling & Software 24 (9) (2009) 1122–1126. doi:10.1016/j.envsoft.2009.02.010.

[103] X. He, L. Ren, A multiscale finite element linearization scheme for the unsaturated flow problems in heterogeneous porous media, Water Resources Research 42 (8). doi:10.1029/2006wr004905.

[104] I. Ginzburg, J.-P. Carlier, C. Kao, Lattice Boltzmann approach to Richards' equation, in: Computational Methods in Water Resources: Volume 1, Elsevier, 2004, pp. 583–595. doi:10.1016/s0167-5648(04)80083-2.

[105] C. E. Kees, C. T. Miller, Higher order time integration methods for two-phase flow, Advances in Water Resources 25 (2) (2002) 159–177. doi:10.1016/s0309-1708(01)00054-9.

[106] W. H. Reed, T. R. Hill, Triangular mesh methods for the neutron transport equation, Proceedings of the American Nuclear Society.

[107] D. N. Arnold, F. Brezzi, B. Cockburn, L. D. Marini, Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems, SIAM Journal on Numerical Analysis 39 (5) (2002) 1749–1779. doi:10.1137/s0036142901384162.

[108] V. Dolejší, M. Kuraz, P. Solin, Adaptive higher-order space-time discontinuous Galerkin method for the computer simulation of variably-saturated porous media flows, Applied Mathematical Modelling 72 (2019) 276–305. doi:10.1016/j.apm.2019.02.037.

[109] S. Sun, M. F. Wheeler, Symmetric and Nonsymmetric Discontinuous Galerkin Methods for Reactive Transport in Porous Media, SIAM Journal on Numerical Analysis 43 (1) (2005) 195–219. doi:10.1137/s003614290241708x.

[110] B. Kane, R. Klöfkorn, C. Gersbacher, hp–Adaptive Discontinuous Galerkin Methods for Porous Media Flow, in: Springer Proceedings in Mathematics & Statistics, Springer International Publishing, 2017, pp. 447–456. doi:10.1007/978-3-319-57394-6_47.

[111] B. Kane, R. Klöfkorn, A. Dedner, Adaptive Discontinuous Galerkin Methods for Flow in Porous Media, in: Lecture Notes in Computational Science and Engineering, Springer International Publishing, 2019, pp. 367–378. doi:10.1007/978-3-319-96415-7_32.

[112] P. Šolín, M. Kuraz, Solving the nonstationary Richards equation with adaptive hp-FEM, Advances in Water Resources 34 (9) (2011) 1062–1081. doi:10.1016/j.advwatres.2011.04.020.

[113] W. F. Mitchell, M. A. McClain, A Comparison of hp-Adaptive Strategies for Elliptic Partial Differential Equations, ACM Transactions on Mathematical Software 41 (1) (2014) 1–39. doi:10.1145/2629459.

[114] H. W. Alt, S. Luckhaus, Quasilinear elliptic-parabolic differential equations, Mathematische Zeitschrift 183 (3) (1983) 311–341. doi:10.1007/bf01176474.

[115] F. Otto, L^1-Contraction and Uniqueness for Quasilinear Elliptic–Parabolic Equations, Journal of Differential Equations 131 (1) (1996) 20–38. doi:10.1006/jdeq.1996.0155.

[116] F. Otto, L^1-Contraction And Uniqueness For Unstationary Saturated-Unsaturated Porous Media Flow, Adv. Math. Sci. Appl. 7 (2) (1997) 537–553.

[117] B. Schweizer, Regularization of outflow problems in unsaturated porous media with dry regions, Journal of Differential Equations 237 (2) (2007) 278–306. doi:10.1016/j.jde.2007.03.011.

[118] I. S. Pop, B. Schweizer, Regularization Schemes For Degenerate Richards Equations And Outflow Conditions, Mathematical Models and Methods in Applied Sciences 21 (08) (2011) 1685–1712. doi:10.1142/s0218202511005532.

[119] Y. Epshteyn, B. Rivière, Estimation of penalty parameters for symmetric interior penalty Galerkin methods, Journal of Computational and Applied Mathematics 206 (2) (2007) 843–872. doi:10.1016/j.cam.2006.08.029.

[120] P. Castillo, Performance of Discontinuous Galerkin Methods for Elliptic PDEs, SIAM Journal on Scientific Computing 24 (2) (2002) 524–547. doi:10.1137/s1064827501388339.

[121] P. van Slingerland, C. Vuik, Fast linear solver for diffusion problems with applications to pressure computation in layered domains, Computational Geosciences 18 (3-4) (2014) 343–356. doi:10.1007/s10596-014-9400-8.

[122] D. Schötzau, L. Zhu, A robust a-posteriori error estimator for discontinuous Galerkin methods for convection–diffusion equations, Applied Numerical Mathematics 59 (9) (2009) 2236–2255. doi:10.1016/j.apnum.2008.12.014.

[123] V. A. Dobrev, R. D. Lazarov, L. T. Zikatanov, Preconditioning of Symmetric Interior Penalty Discontinuous Galerkin FEM for Elliptic Problems, in: Lecture Notes in Computational Science and Engineering, Springer Berlin Heidelberg, 2008, pp. 33–44. doi:10.1007/978-3-540-75199-1_3.

[124] T. D. Maet, E. Hanert, M. Vanclooster, A fully-explicit discontinuous Galerkin hydrodynamic model for variably-saturated porous media, Journal of Hydrodynamics 26 (4) (2014) 594–607. doi:10.1016/s1001-6058(14)60067-6.

[125] O. Østerby, Five Ways of Reducing the Crank–Nicolson Oscillations, BIT Numerical Mathematics 43 (4) (2003) 811–822. doi:10.1023/b:bitn.0000009942.00540.94.

[126] E. Süli, D. F. Mayers, An Introduction to Numerical Analysis, Cambridge University Press, 2003. doi:10.1017/cbo9780511801181.

[127] V. Dolejší, M. Vlasák, Analysis of a BDF–DGFE scheme for nonlinear convection–diffusion problems, Numerische Mathematik 110 (4) (2008) 405–447. doi:10.1007/s00211-008-0178-2.

[128] A. Hay, S. Etienne, D. Pelletier, A. Garon, hp-Adaptive time integration based on the BDF for viscous flows, Journal of Computational Physics 291 (2015) 151–176. doi:10.1016/j.jcp.2015.03.022.

[129] G. G. Dahlquist, A special stability problem for linear multistep methods, BIT 3 (1) (1963) 27–43. doi:10.1007/bf01963532.

[130] S. Lortie, Simulation d'écoulements á surface libre entre un liquide et un gaz compressible, Mémoire de maîtrise, École Polytechnique de Montréal (2018).
URL https://publications.polymtl.ca/3272/

[131] R. D. Grigorieff, Stability of multistep-methods on variable grids, Numerische Mathematik 42 (3) (1983) 359–377. doi:10.1007/bf01389580.

[132] S. Skelboe, The control of order and steplength for backward differentiation methods, BIT 17 (1) (1977) 91–107. doi:10.1007/bf01932401.

[133] K. Rathfelder, L. M. Abriola, Mass conservative numerical solutions of the head-based Richards equation, Water Resources Research 30 (9) (1994) 2579–2586. doi:10.1029/94wr01302.

[134] K. Huang, B. P. Mohanty, M. T. van Genuchten, A new convergence criterion for the modified Picard iteration method to solve the variably saturated flow equation, Journal of Hydrology 178 (1-4) (1996) 69–91. doi:10.1016/0022-1694(95)02799-8.

[135] L. Bergamaschi, M. Putti, Mixed finite elements and Newton-type linearizations for the solution of Richards' equation, International Journal for Numerical Methods in Engineering 45 (8) (1999) 1025–1046. doi:10.1002/(sici)1097-0207(19990720)45:8<1025::aid-nme615>3.0.co;2-g.

[136] J. Šimůnek, M. Šejna, H. Saito, M. Sakai, M. T. van Genuchten, The HYDRUS-1D Software Package for Simulating the Movement of Water, Heat, and MultipleSolutes in Variably Saturated Media, Version 4.08, Department of Environmental Sciences, University of California Riverside, 2009.

[137] D. Kavetski, P. Binning, S. W. Sloan, Adaptive time stepping and error control in a mass conservative numerical solution of the mixed form of Richards equation, Advances in Water Resources 24 (6) (2001) 595–605. doi:10.1016/s0309-1708(00)00076-2.

[138] D. Kavetski, P. Binning, S. W. Sloan, Noniterative time stepping schemes with adaptive truncation error control for the solution of Richards equation, Water Resources Research 38 (10) (2002) 29–1–29–10. doi:10.1029/2001wr000720.

[139] F. H. Maina, P. Ackerer, Ross scheme, Newton–Raphson iterative methods and time-stepping strategies for solving the mixed form of Richards' equation, Hydrology and Earth System Sciences 21 (6) (2017) 2667–2683. doi:10.5194/hess-21-2667-2017.

[140] G. Dhatt, G. Touzot, Une présentation de la méthode des éléments finis, Maloine S.A. Éditeur, 1981.

[141] Y. Notay, An aggregation-based algebraic multigrid method., ETNA. Electronic Transactions on Numerical Analysis [electronic only] 37 (2010) 123–146.
URL http://eudml.org/doc/230391

[142] Y. Notay, Aggregation-Based Algebraic Multigrid for Convection-Diffusion Equations, SIAM Journal on Scientific Computing 34 (4) (2012) A2288–A2316. doi:10.1137/110835347.

[143] A. Napov, Y. Notay, An Algebraic Multigrid Method with Guaranteed Convergence Rate, SIAM Journal on Scientific Computing 34 (2) (2012) A1079–A1109. doi:10.1137/100818509.

[144] D. A. Di Pietro, A. Ern, Mathematical Aspects of Discontinuous Galerkin Methods, Springer Berlin Heidelberg, 2012. doi:10.1007/978-3-642-22980-0.

[145] E. Cuthill, J. McKee, Reducing the bandwidth of sparse symmetric matrices, in: Proceedings of the 1969 24th national conference on -, ACM Press, 1969. doi:10.1145/800195.805928.

[146] F. Alauzet, A. Loseille, On the Use of Space Filling Curves for Parallel Anisotropic Mesh Adaptation, in: Proceedings of the 18th International Meshing Roundtable, Springer Berlin Heidelberg, 2009, pp. 337–357. doi:10.1007/978-3-642-04319-2_20.

[147] C. Johnson, Adaptive finite element methods for diffusion and convection problems, Computer Methods in Applied Mechanics and Engineering 82 (1-3) (1990) 301–322. doi:10.1016/0045-7825(90)90169-m.

[148] I. Babuška, W. C. Rheinboldt, Error Estimates for Adaptive Finite Element Computations, SIAM Journal on Numerical Analysis 15 (4) (1978) 736–754.

[149] M. Ainsworth, J. T. Oden, A Posteriori Error Estimation in Finite Element Analysis, Wiley, New York, 2000.

[150] M. Ainsworth, J. T. Oden, A posteriori error estimation in finite element analysis, Computer Methods in Applied Mechanics and Engineering 142 (1-2) (1997) 1–88. doi:10.1016/s0045-7825(96)01107-3.

[151] R. Verfürth, A Posteriori Error Estimation Techniques for Finite Element Methods, Oxford University Press, 2013. doi:10.1093/acprof:oso/9780199679423.001.0001.

[152] L. Demkowicz, J. T. Oden, W. Rachowicz, O. Hardy, Toward a universal adaptive finite element strategy, part 1. constrained approximation and data structure, Computer Methods in Applied Mechanics and Engineering 77 (1-2) (1989) 79–112. doi:10.1016/0045-7825(89)90129-1.

[153] J. T. Oden, L. Demkowicz, W. Rachowicz, T. A. Westermann, Toward a universal adaptive finite element strategy, part 2. a posteriori error estimation, Computer Methods in Applied Mechanics and Engineering 77 (1-2) (1989) 113–180. doi:10.1016/0045-7825(89)90130-8.

[154] W. Rachowicz, J. T. Oden, L. Demkowicz, Toward a universal adaptive finite element strategy part 3. design of meshes, Computer Methods in Applied Mechanics and Engineering 77 (1-2) (1989) 181–212. doi:10.1016/0045-7825(89)90131-x.

[155] P. Šolín, K. Segeth, I. Doležel, Higher-Order Finite Element Methods, Chapman & Hall/CRC, Boca Raton, FL, 2003.

[156] P. Šolín, L. Demkowicz, Goal-oriented hp-adaptivity for elliptic problems, Computer Methods in Applied Mechanics and Engineering 193 (6-8) (2004) 449–468. doi:10.1016/j.cma.2003.09.015.

[157] M. Ainsworth, J. T. Oden, A unified approach to a posteriori error estimation using element residual methods, Numerische Mathematik 65 (1) (1993) 23–50. doi:10.1007/bf01385738.

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

[159] R. Becker, R. Rannacher, An optimal control approach to a posteriori error estimation in finite element methods, Acta Numerica 10 (2001) 1–102. doi:10.1017/s0962492901000010.

[160] M. J. Berger, P. Colella, Local adaptive mesh refinement for shock hydrodynamics, Journal of Computational Physics 82 (1) (1989) 64–84. doi:10.1016/0021-9991(89)90035-1.

[161] K. F. C. Yiu, D. M. Greaves, S. Cruz, A. Saalehi, A. G. L. Borthwick, Quadtree grid generation: Information handling, boundary fitting and CFD applications, Computers & Fluids 25 (8) (1996) 759–769. doi:10.1016/s0045-7930(96)00029-1.

[162] C. Min, F. Gibou, A second order accurate level set method on non-graded adaptive cartesian grids, Journal of Computational Physics 225 (1) (2007) 300–321. doi:10.1016/j.jcp.2006.11.034.

[163] M. Zhang, W. M. Wu, A two dimensional hydrodynamic and sediment transport model for dam break based on finite volume method with quadtree grid, Applied Ocean Research 33 (4) (2011) 297–308. doi:10.1016/j.apor.2011.07.004.

[164] F. Losasso, F. Gibou, R. Fedkiw, Simulating water and smoke with an octree data structure, ACM Transactions on Graphics 23 (3) (2004) 457. doi:10.1145/1015706.1015745.

[165] D. Fuster, G. Agbaglah, C. Josserand, S. Popinet, S. Zaleski, Numerical simulation of droplets, bubbles and waves: state of the art, Fluid Dynamics Research 41 (6) (2009) 065001. doi:10.1088/0169-5983/41/6/065001.

[166] T. Coupez, E. Hachem, Solution of high-Reynolds incompressible flow with stabilized finite element and adaptive anisotropic meshing, Computer Methods in Applied Mechanics and Engineering 267 (2013) 65–85. doi:10.1016/j.cma.2013.08.004.

[167] E. Hachem, S. Feghali, R. Codina, T. Coupez, Immersed stress method for fluid-structure interaction using anisotropic mesh adaptation, International Journal for Numerical Methods in Engineering 94 (9) (2013) 805–825. doi:10.1002/nme.4481.

[168] C. Bahbah, M. Khalloufi, A. Larcher, Y. Mesri, T. Coupez, R. Valette, E. Hachem, Conservative and adaptive level-set method for the simulation of two-fluid flows, Computers & Fluids 191 (2019) 104223. doi:10.1016/j.compfluid.2019.06.022.

[169] H. Dong, Z. Qiao, S. Sun, T. Tang, Adaptive moving grid methods for two-phase flow in porous media, Journal of Computational and Applied Mathematics 265 (2014) 139–150. doi:10.1016/j.cam.2013.09.027.

[170] W. Huang, X. Zhan, Adaptive moving mesh modeling for two dimensional groundwater flow and transport (2005). doi:10.1090/conm/383/07168.

[171] J. Z. X. Zheng, C. P. T. Groth, Block-Based Adaptive Mesh Renement Finite-Volume Scheme forHybrid Multi-Block Meshes, in: Seventh International Conference on Computational Fluid Dynamics (ICCFD7), Big Island, Hawaii, 2012.

[172] A. Giuliani, L. Krivodonova, Adaptive mesh refinement on graphics processing units for applications in gas dynamics, Journal of Computational Physics 381 (2019) 67–90. doi:10.1016/j.jcp.2018.12.019.

[173] D. Dunning, W. Marts, R. W. Robey, P. Bridges, Adaptive mesh refinement in the fast lane, Journal of Computational Physics 406 (2020) 109193. doi:10.1016/j.jcp.2019.109193.

[174] T. Altazin, M. Ersoy, F. Golay, D. Sous, L. Yushchenko, Numerical investigation of BB-AMR scheme using entropy production as refinement criterion, International Journal of Computational Fluid Dynamics 30 (3) (2016) 256–271. doi:10.1080/10618562.2016.1194977.

[175] M. Kuraz, P. Mayer, P. Pech, Solving the nonlinear Richards equation model with adaptive domain decomposition, Journal of Computational and Applied Mathematics 270 (2014) 2–11. doi:10.1016/j.cam.2014.03.010.

[176] K. Pons, M. Ersoy, Adaptive mesh refinement method. Part 1: Automatic thresholding based on a distribution function, working paper or preprint (Jul. 2019).
URL https://hal.archives-ouvertes.fr/hal-01330679

[177] M. Ersoy, F. Golay, L. Yushchenko, Adaptive multiscale scheme based on numerical density of entropy production for conservation laws, Open Mathematics 11 (8). doi:10.2478/s11533-013-0252-6.

[178] A. Ern, A. F. Stephansen, M. Vohralík, Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection–diffusion–reaction problems, Journal of Computational and Applied Mathematics 234 (1) (2010) 114–130. doi:10.1016/j.cam.2009.12.009.

[179] R. Verfürth, A posteriori error estimates for nonlinear problems: $L^r(0,T;W^{1,\rho}(\Omega))$-error estimates for finite element discretizations of parabolic equations, Numerical Methods for Partial Differential Equations 14 (4) (1998) 487–518. doi:10.1002/(SICI)1098-2426(199807)14:4<487::AID-NUM4>3.0.CO;2-G.

[180] R. Verfürth, A posteriori error estimates for finite element discretizations of the heat equation, CALCOLO 40 (3) (2003) 195–212. doi:10.1007/s10092-003-0073-2.

[181] R. Verfürth, A posteriori error estimatesfor non-linear parabolic equations, Tech. Rep. report 361, Fakultät für mathematik, Ruhr-Universität, Bochum, D-44780 Bochun (Dec. 2004).

[182] J. M. Melenk, B. I. Wohlmuth, On residual-based a posteriori error estimation in hp-FEM, Advances in Computational Mathematics 15 (1/4) (2001) 311–331. doi:10.1023/a:1014268310921.

[183] P. Houston, D. Schötzau, T. P. Wihler, Energy Norm A Posteriori Error Estimation of hp-Adaptive Discontinuous Galerkin Methods for Elliptic Problems, Mathematical Models and Methods in Applied Sciences 17 (01) (2007) 33–62. doi:10.1142/s0218202507001826.

[184] L. Zhu, D. Schotzau, A robust a posteriori error estimate for hp-adaptive DG methods for convection-diffusion equations, IMA Journal of Numerical Analysis 31 (3) (2010) 971–1005. doi:10.1093/imanum/drp038.

[185] A. Ern, J. Proft, Multi-algorithmic methods for coupled hyperbolic-parabolic problems, International Journal of Numerical Analysis and Modeling 3 (1) (2006) 94–114. doi:10.1016/j.aml.2004.05.019.

[186] P. Houston, C. Schwab, E. Süli, Discontinuoushp-Finite Element Methods for Advection-Diffusion-Reaction Problems, SIAM Journal on Numerical Analysis 39 (6) (2002) 2133–2163. doi:10.1137/s0036142900374111.

[187] J.-P. Croisille, A. Ern, T. Lelièvre, J. Proft, Analysis and simulation of a coupled hyperbolic/parabolic model problem, Journal of Numerical Mathematics 13 (2). doi:10.1515/1569395054012776.

[188] E. Burman, P. Zunino, A Domain Decomposition Method Based on Weighted Interior Penalties for Advection-Diffusion-Reaction Problems, SIAM Journal on Numerical Analysis 44 (4) (2006) 1612–1638. doi:10.1137/050634736.

[189] K. Shahbazi, An explicit expression for the penalty parameter of the interior penalty method, Journal of Computational Physics 205 (2) (2005) 401–407. doi:10.1016/j.jcp.2004.11.017.

[190] R. Hartmann, P. Houston, An optimal order interior penalty discontinuous Galerkin discretization of the compressible Navier–Stokes equations, Journal of Computational Physics 227 (22) (2008) 9670–9685. doi:10.1016/j.jcp.2008.07.015.

[191] D. Wirasaet, S. Tanaka, E. J. Kubatko, J. J. Westerink, C. Dawson, A performance comparison of nodal discontinuous Galerkin methods on triangles and quadrilaterals, International Journal for Numerical Methods in Fluids 64 (10-12) (2010) 1336–1362. doi:10.1002/fld.2376.

[192] R. Haverkamp, M. Vauclin, J. Touma, P. J. Wierenga, G. Vachaud, A Comparison of Numerical Simulation Models For One-Dimensional Infiltration, Soil Science Society of America Journal 41 (2) (1977) 285. doi:10.2136/sssaj1977.03615995004100020024x.

[193] J. R. Philip, The theory of infiltration: 1. the infiltration equation and its solution, Soil Science 83 (5) (1957) 345–358.
URL https://journals.lww.com/soilsci/Fulltext/1957/05000/THE_THEORY_OF_INFILTRATION__1__THE_INFILTRATION.2.aspx

[194] C. W. Li, A simplified Newton Iteration Method with linear finite elements for transient unsaturated flow, Water Resources Research 29 (4) (1993) 965–971. doi:10.1029/92wr02891.

[195] A. I. El-Kadi, G. Ling, The Courant and Peclet Number criteria for the numerical solution of the Richards Equation, Water Resources Research 29 (10) (1993) 3485–3494. doi:10.1029/93wr00929.

[196] D. J. Polmann, E. G. Vomvoris, D. McLaughlin, E. M. Hammick, L. W. Gelhar, Application of Stochastic Methods to the Simulation of Large-scale Unsaturated Flow and Transport (1988).

[197] F. T. Tracy, Clean two- and three-dimensional analytical solutions of Richards' equation for testing numerical solvers, Water Resources Research 42 (8). doi:10.1029/2005wr004638.

[198] F. T. Tracy, Three-dimensional analytical solutions of Richards' equation for a box-shaped soil sample with piecewise-constant head boundary conditions on the top, Journal of Hydrology 336 (3-4) (2007) 391–400. doi:10.1016/j.jhydrol.2007.01.011.

[199] F. T. Tracy, Testing Computational Algorithms for Unsaturated Flow, The Open Hydrology Journal 4 (1) (2010) 227–235. doi:10.2174/1874378101004010227.

[200] G. Vachaud, M. Vauclin, D. Khanji, Étude expérimentale des transferts bidimensionnels dans la zone non saturée application à l'étude du drainage d'une nappe à surface libre, La Houille Blanche (1) (1973) 65–74. doi:10.1051/lhb/1973005.

[201] M. Vauclin, G. Vachaud, J. Khanji, Two dimensional numerical analysis of transient water transfer in saturated-unsaturated soils, in: G. C. Vansteenkiste (Ed.), Computer Simulation of Water Resources Systems, IFIP Work. Conf., Amsterdam, North-holland Publishing company, 1975, pp. 299–323, Modeling and Simulation of Water Resources Systems.

[202] G. Vachaud, M. Vauclin, Comments on 'A numerical model based on coupled one-dimensional Richards and Boussinesq equations' by Mary F. Pikul, Robert L. Street, and Irwin Remson, Water Resources Research 11 (3) (1975) 506–509. doi:10.1029/wr011i003p00506.

[203] M. Vauclin, D. Khanji, G. Vachaud, Experimental and numerical study of a transient, two-dimensional unsaturated-saturated water table recharge problem, Water Resources Research 15 (5) (1979) 1089–1101. doi:10.1029/wr015i005p01089.

[204] T. P. Clement, W. R. Wise, F. J. Molz, A physically based, two-dimensional, finite-difference algorithm for modeling variably saturated flow, Journal of Hydrology 161 (1-4) (1994) 71–90. doi:10.1016/0022-1694(94)90121-x.

[205] C. Shen, M. S. Phanikumar, A process-based, distributed hydrologic model based on a large-scale method for surface–subsurface coupling, Advances in Water Resources 33 (12) (2010) 1524–1541. doi:10.1016/j.advwatres.2010.09.002.

[206] N. K. C. Twarakavi, J. Šimůnek, S. Seo, Evaluating Interactions between Groundwater and Vadose Zone Using the HYDRUS-Based Flow Package for MODFLOW, Vadose Zone Journal 7 (2) (2008) 757–768. doi:10.2136/vzj2007.0082.

[207] X. Xu, G. Huang, H. Zhan, Z. Qu, Q. Huang, Integration of SWAP and MODFLOW-2000 for modeling groundwater dynamics in shallow water table areas, Journal of Hydrology 412-413 (2012) 170–181. doi:10.1016/j.jhydrol.2011.07.002.

[208] Y. Zha, L. Shi, M. Ye, J. Yang, A generalized Ross method for two- and three-dimensional variably saturated flow, Advances in Water Resources 54 (2013) 67–77. doi:10.1016/j.advwatres.2013.01.002.

[209] J. M. Fleureau, J. J. Fry, Validation des modèles de couplage sur ouvrages réels, Tech. rep., Ministère de la Recherche et de la Technologie/CNRS. Rapport Scientifique du GRECO "Géomatériaux" (1991).

[210] P. S. Huyakorn, G. F. Pinder, The Computational Methods in Subsurface Flow, Academic Press, 1983. doi:10.1016/c2012-0-01564-5.

[211] G. F. Pinder, W. G. Gray, Finite Element Simulation in Surface and Subsurface Hydrology, Academic Press, New York, 1977.

[212] L. Yao, D. Wang, M. Hooshyar, A. Singh, M. Sivapalan, Time Compression Approximation Relationship for Infiltration in the Presence of a Shallow Water Table: Evaluating the Role of Péclet Number, Water Resources Research 54 (11) (2018) 9384–9397. doi:10.1029/2018wr023293.

[213] J. R. Philip, Theory of Infiltration, in: Advances in Hydroscience, Elsevier, 1969, pp. 215–296. doi:10.1016/b978-1-4831-9936-8.50010-6.

[214] S. Malott, D. M. O'Carroll, C. E. Robinson, Dynamic groundwater flows and geochemistry in a sandy nearshore aquifer over a wave event, Water Resources Research 52 (7) (2016) 5248–5264. doi:10.1002/2015wr017537.

[215] L. D. Wright, A. D. Short, Morphodynamic variability of surf zones and beaches: A synthesis, Marine Geology 56 (1-4) (1984) 93–118. doi:10.1016/0025-3227(84)90008-2.

[216] C. Robinson, B. Gibbes, L. Li, Driving mechanisms for groundwater flow and salt transport in a subterranean estuary, Geophysical Research Letters 33 (3). doi:10.1029/2005gl025247.

[217] W. C. Burnett, P. K. Aggarwal, A. Aureli, H. Bokuniewicz, J. E. Cable, M. A. Charette, E. Kontar, S. Krupa, K. M. Kulkarni, A. Loveless, W. S. Moore, J. A. Oberdorfer, J. Oliveira, N. Ozyurt, P. Povinec, A. M. G. Privitera, R. Rajar, R. T. Ramessur, J. Scholten, T. Stieglitz, M. Taniguchi, J. V. Turner, Quantifying submarine groundwater discharge in the coastal zone via multiple methods, Science of The Total Environment 367 (2-3) (2006) 498–543. doi:10.1016/j.scitotenv.2006.05.009.

[218] P. Anschutz, T. Smith, A. Mouret, J. Deborde, S. Bujan, D. Poirier, P. Lecroart, Tidal sands as biogeochemical reactors, Estuarine, Coastal and Shelf Science 84 (1) (2009) 84–90. doi:10.1016/j.ecss.2009.06.015.

[219] A. H. Sawyer, O. Lazareva, K. D. Kroeger, K. Crespo, C. S. Chan, T. Stieglitz, H. A. Michael, Stratigraphic controls on fluid and solute fluxes across the sediment-water interface of an estuary, Limnology and Oceanography 59 (3) (2014) 997–1010. doi:10.4319/lo.2014.59.3.0997.

[220] C. Charbonnier, P. Anschutz, D. Poirier, S. Bujan, P. Lecroart, Aerobic respiration in a high-energy sandy beach, Marine Chemistry 155 (2013) 10–21. doi:10.1016/j.marchem.2013.05.003.

[221] S. M. McAllister, J. M. Barnett, J. W. Heiss, A. J. Findlay, D. J. MacDonald, C. L. Dow, G. W. Luther, H. A. Michael, C. S. Chan, Dynamic hydrologic and biogeochemical processes drive microbially enhanced iron and sulfur cycling within the intertidal mixing zone of a beach aquifer, Limnology and Oceanography 60 (1) (2015) 329–345. doi:10.1002/lno.10029.

[222] J. W. Heiss, V. E. A. Post, T. Laattoe, C. J. Russoniello, H. A. Michael, Physical Controls on Biogeochemical Processes in Intertidal Zones of Beach Aquifers, Water Resources Research 53 (11) (2017) 9225–9244. doi:10.1002/2017wr021110.

[223] P. Nielsen, Tidal dynamics of the water table in beaches, Water Resources Research 26 (9) (1990) 2127–2134. doi:10.1029/wr026i009p02127.

[224] X. Bertin, A. de Bakker, A. van Dongeren, G. Coco, G. André, F. Ardhuin, P. Bonneton, F. Bouchette, B. Castelle, W. C. Crawford, M. Davidson, M. Deen, G. Dodet, T. Guérin, K. Inch, F. Leckler, R. McCall, H. Muller, M. Olabarrieta, D. Roelvink, G. Ruessink, D. Sous, É. Stutzmann, M. Tissier, Infragravity waves: From driving mechanisms to impacts, Earth-Science Reviews 177 (2018) 774–799. doi:10.1016/j.earscirev.2018.01.002.

[225] X. Chen, Y.-M. Chiew, Velocity Distribution of Turbulent Open-Channel Flow with Bed Suction, Journal of Hydraulic Engineering 130 (2) (2004) 140–148. doi:10.1061/(asce)0733-9429(2004)130:2(140).

[226] D. C. Conley, D. L. Inman, Ventilated oscillatory boundary layers, Journal of Fluid Mechanics 273 (1994) 261–284. doi:10.1017/s002211209400193x.

[227] S. Corvaro, E. Seta, A. Mancinelli, M. Brocchini, Flow dynamics on a porous medium, Coastal Engineering 91 (2014) 280–298. doi:10.1016/j.coastaleng.2014.06.001.

[228] S. Corvaro, M. Miozzi, M. Postacchini, A. Mancinelli, M. Brocchini, Fluid–particle interaction and generation of coherent structures over permeable beds: an experimental analysis, Advances in Water Resources 72 (2014) 97–109. doi:10.1016/j.advwatres.2014.05.015.

[229] I. L. Turner, G. Masselink, Swash infiltration-exfiltration and sediment transport, Journal of Geophysical Research: Oceans 103 (C13) (1998) 30813–30824. doi:10.1029/98jc02606.

[230] T. Butt, P. Russell, I. Turner, The influence of swash infiltration–exfiltration on beach face sediment transport: onshore or offshore?, Coastal Engineering 42 (1) (2001) 35–52. doi:10.1016/s0378-3839(00)00046-6.

[231] G. Masselink, L. Li, The role of swash infiltration in determining the beachface gradient: a numerical study, Marine Geology 176 (1-4) (2001) 139–156. doi:10.1016/s0025-3227(01)00161-x.

[232] T. V. Karambas, Modelling of Infiltration-Exfiltration Effects of Cross-Shore Sediment Transport in the Swash Zone, Coastal Engineering Journal 45 (1) (2003) 63–82. doi:10.1142/s057856340300066x.

[233] M. J. Austin, G. Masselink, Swash–groundwater interaction on a steep gravel beach, Continental Shelf Research 26 (20) (2006) 2503–2519. doi:10.1016/j.csr.2006.07.031.

[234] R. Bakhtyar, A. Brovelli, D. A. Barry, L. Li, Wave-induced water table fluctuations, sediment transport and beach profile change: Modeling and comparison with large-scale laboratory experiments, Coastal Engineering 58 (1) (2011) 103–118. doi:10.1016/j.coastaleng.2010.08.004.

[235] G. Masselink, A. Ruju, D. Conley, I. Turner, G. Ruessink, A. Matias, C. Thompson, B. Castelle, J. Puleo, V. Citerone, G. Wolters, Large-scale Barrier Dynamics Experiment II (BARDEX II): Experimental design, instrumentation, test program, and data set, Coastal Engineering 113 (2016) 3–18. doi:10.1016/j.coastaleng.2015.07.009.

[236] P. Chardón-Maldonado, J. C. Pintado-Patiño, J. A. Puleo, Advances in swash-zone research: Small-scale hydrodynamic and sediment transport processes, Coastal Engineering 115 (2016) 8–25. doi:10.1016/j.coastaleng.2015.10.008.

[237] E. Perera, F. Zhu, N. Dodd, R. Briganti, C. Blenkinsopp, I. L. Turner, Surface-groundwater flow numerical model for barrier beach with exfiltration incorporated bottom boundary layer model, Coastal Engineering 146 (2019) 47–64. doi:10.1016/j.coastaleng.2018.12.002.

[238] D. Sous, A. Lambert, V. Rey, H. Michallet, Swash–groundwater dynamics in a sandy beach laboratory experiment, Coastal Engineering 80 (2013) 122–136. doi:10.1016/j.coastaleng.2013.05.006.

[239] D. P. Horn, Measurements and modelling of beach groundwater flow in the swash-zone: a review, Continental Shelf Research 26 (5) (2006) 622–652. doi:10.1016/j.csr.2006.02.001.

[240] I. L. Turner, S. P. Leatherman, Beach Dewatering as a 'Soft' Engineering Solution to Coastal Erosion-A History and Critical Review, Journal of Coastal Research 13 (4) (1997) 1050–1063.

[241] H.-Y. Kang, P. Nielsen, Watertable Dynamics in Coastal Areas, in: Coastal Engineering 1996, American Society of Civil Engineers, 1997. doi:10.1061/9780784402429.358.

[242] T. V. Karambas, Prediction of sediment transport in the swash-zone by using a nonlinear wave model, Continental Shelf Research 26 (5) (2006) 599–609. doi:10.1016/j.csr.2006.01.014.

[243] L. Li, D. A. Barry, C. B. Pattiaratchi, Numerical modelling of tide-induced beach water table fluctuations, Coastal Engineering 30 (1-2) (1997) 105–123. doi:10.1016/s0378-3839(96)00038-5.

[244] M. R. A. van Gent, The modelling of wave action on and in coastal structures, Coastal Engineering 22 (3-4) (1994) 311–339. doi:10.1016/0378-3839(94)90041-8.

[245] K. Steenhauer, D. Pokrajac, T. O'Donoghue, Numerical model of swash motion and air entrapment within coarse-grained beaches, Coastal Engineering 64 (2012) 113–126. doi:10.1016/j.coastaleng.2012.01.004.

[246] J. Desombre, D. Morichon, M. Mory, SIMULTANEOUS SURFACE AND SUBSURFACE AIR AND WATER FLOWS MODELLING IN THE SWASH ZONE, Coastal Engineering Proceedings 1 (33) (2012) 56. doi:10.9753/icce.v33.currents.56.

[247] R. T. McCall, G. Masselink, D. Roelvink, P. Russell, M. Davidson, T. Poate, MODELLING OVERWASH AND INFILTRATION ON GRAVEL BARRIERS, Coastal Engineering Proceedings 1 (33) (2012) 34. doi:10.9753/icce.v33.currents.34.

[248] R. T. McCall, G. Masselink, T. G. Poate, J. A. Roelvink, L. P. Almeida, M. Davidson, P. E. Russell, Modelling storm hydrodynamics on gravel beaches with XBeach-G, Coastal Engineering 91 (2014) 231–250. doi:10.1016/j.coastaleng.2014.06.007.

[249] S. Weill, E. Mouche, J. Patin, A generalized Richards equation for surface/subsurface flow modelling, Journal of Hydrology 366 (1-4) (2009) 9–20. doi:10.1016/j.jhydrol.2008.12.007.

[250] Y. Pachepsky, D. Timlin, W. Rawls, Generalized Richards' equation to simulate water transport in unsaturated soils, Journal of Hydrology 272 (1-4) (2003) 3–13. doi:10.1016/s0022-1694(02)00251-2.

[251] I. L. Turner, P. Nielsen, Rapid water table fluctuations within the beach face: Implications for swash zone sediment mobility?, Coastal Engineering 32 (1) (1997) 45–59. doi:10.1016/s0378-3839(97)00015-x.

[252] A. Furman, Modeling Coupled Surface-Subsurface Flow Processes: A Review, Vadose Zone Journal 7 (2) (2008) 741–756. doi:10.2136/vzj2007.0065.

[253] M. Zijlema, G. Stelling, P. Smit, SWASH: An operational public domain code for simulating wave fields and rapidly varied flows in coastal waters, Coastal Engineering 58 (10) (2011) 992–1012. doi:10.1016/j.coastaleng.2011.05.015.

[254] G. A. Kikkert, D. Pokrajac, T. O'Donoghue, K. Steenhauer, Experimental study of bore-driven swash hydrodynamics on permeable rough slopes, Coastal Engineering 79 (2013) 42–56. doi:10.1016/j.coastaleng.2013.04.008.

[255] G. A. Kikkert, T. O'Donoghue, D. Pokrajac, N. Dodd, Experimental study of bore-driven swash hydrodynamics on impermeable rough slopes, Coastal Engineering 60 (2012) 149–166. doi:10.1016/j.coastaleng.2011.09.006.

[256] J.-B. Clément, D. Sous, F. Golay, M. Ersoy, Wave-driven Groundwater Flows in Sandy Beaches: A Richards Equation-based Model, Journal of Coastal Research 95 (sp1) (2020) 1047. doi:10.2112/si95-204.1.

[257] J.-B. Clément, F. Golay, M. Ersoy, D. Sous, Adaptive Discontinuous Galerkin Method for Richards Equation, in: Topical Problems of Fluid Mechanics 2020, Institute of Thermomechanics, AS CR, v.v.i., 2020. doi:10.14311/tpfm.2020.004.

[258] J.-B. Clément, M. Ersoy, F. Golay, D. Sous, Discontinuous Galerkin Method for Steady-State Richards Equation, in: Topical Problems of Fluid Mechanics 2019, Institute of Thermomechanics, AS CR, v.v.i., 2019. doi:10.14311/tpfm.2019.008.

Jean-Baptiste CLÉMENT

Simulation numérique des écoulements en milieu poreux non-saturés par une méthode de Galerkine discontinue adaptative : application aux plages sableuses

Résumé : Les écoulements en milieux poreux non-saturés sont modélisés par l'équation de Richards qui est une équation non-linéaire parabolique dégénérée. Ses limites et les défis que soulèvent sa résolution numérique sont présentés. L'obtention de résultats robustes, précis et efficaces est difficile en particulier à cause des fronts de saturation raides et dynamiques induits par les propriétés hydrauliques non-linéaires. L'équation de Richards est discrétisée par une méthode de Galerkine discontinue en espace et des formules de différentiation rétrograde en temps. Le schéma numérique résultant est conservatif, d'ordre élevé et très flexible. Ainsi, des conditions aux limites complexes sont facilement intégrées comme la condition de suintement ou un forçage dynamique. De plus, une stratégie adaptative est proposée. Un pas de temps adaptatif rend la convergence non-linéaire robuste et un raffinement de maillage adaptatif basée sur des blocs est utilisée pour atteindre la précision requise efficacement. Un indicateur d'erreur *a posteriori* approprié aide le maillage à capturer les fronts de saturation raides qui sont également mieux approximés par une discontinuité introduite dans la solution grâce à une méthode de Galerkine discontinue pondérée. L'approche est validée par divers cas-tests et un benchmark 2D. Les simulations numériques sont comparées à des expériences de laboratoire de recharge/drainage de nappe et une expérience à grande échelle d'humidification, suite à la mise en eau du barrage multi-matériaux de La Verne. Ce cas exigeant montre les potentialités de la stratégie développée dans cette thèse. Enfin, des applications sont menées pour simuler les écoulements souterrains sous la zone de jet de rive de plages sableuses en comparaison avec des observations expérimentales.

Mots-clés : Équation de Richards, suintement, méthode GD pondérée, raffinement de maillage, estimation d'erreur *a posteriori*, BDF, hydrodynamique souterraine du jet de rive.

Numerical simulation of flows in unsaturated porous media by an adaptive discontinuous Galerkin method: application to sandy beaches

Abstract: Flows in unsaturated porous media are modelled by the Richards' equation which is a degenerate parabolic nonlinear equation. Its limitations and the challenges raised by its numerical solution are laid out. Getting robust, accurate and cost-effective results is difficult in particular because of moving sharp wetting fronts due to the nonlinear hydraulic properties. Richards' equation is discretized by a discontinuous Galerkin method in space and backward differentiation formulas in time. The resulting numerical scheme is conservative, high-order and very flexible. Thereby, complex boundary conditions are included easily such as seepage condition or dynamic forcing. Moreover, an adaptive strategy is proposed. Adaptive time stepping makes nonlinear convergence robust and a block-based adaptive mesh refinement is used to reach required accuracy cost-effectively. A suitable *a posteriori* error indicator helps the mesh to capture sharp wetting fronts which are also better approximated by a discontinuity introduced in the solution thanks to a weighted discontinuous Galerkin method. The approach is checked through various test-cases and a 2D benchmark. Numerical simulations are compared with laboratory experiments of water table recharge/drainage and a largescale experiment of wetting, following reservoir impoundment of the multi-materials La Verne dam. This demanding case shows the potentiality of the strategy developed in this thesis. Finally, applications are handled to simulate groundwater flows under the swash zone of sandy beaches in comparison with experimental observations.

Keywords: Richards' equation, seepage, weighted DG method, AMR, *a posteriori* error estimation, BDF, swash groundwater.

Université de Toulon