Mass-Conservative Finite Volume Methods for the Richards’ equation

Lecturer: G. Manzini

$^{1}$IMATI - CNR, Pavia

Siena 2006

(Joint collaboration with Stefano Ferraris, Univ. of Turin, Italy)
Richards’ equation: mixed $\theta$–$\psi$ formulation

- The solution $\psi(x, t)$ satisfies
  \[
  \frac{\partial \theta(\psi)}{\partial t} - \text{div}\left[ K(\psi) \nabla (\psi + z) \right] = s, \quad \text{in} \quad \Omega \times \mathbb{R}^+,
  \]

- $\Omega \subset \mathbb{R}^d$, computational domain
- $\psi$, hydraulic pressure head;
- $\theta(\psi)$, volumetric water content;
- $K(\psi)$, hydraulic conductivity tensor;
- $z$, vertical coordinate (positive upward)
- $s$, production/consumption source term;
Richards’ equation: head formulation

- **The solution** \( \psi(x, t) \) **satisfies**

\[
\eta(\psi) \frac{\partial \psi}{\partial t} - \text{div} \left[ K(\psi) \nabla (\psi + z) \right] = s, \quad \text{in} \quad \Omega \times \mathbb{R}^+, \\
\]

where

- \( \eta(\psi) = \frac{\partial \theta(\psi)}{\partial \psi} \) **is the general storage term.**

- **Remark:** the two formulations are theoretically equivalent, but leads to different numerical discretizations.
To complete the model... 

...we must assign:

- a proper set of **boundary conditions**:
  
  Dirichlet BCs: \( \psi = g^D \), on \( \Gamma^D \times \mathbb{R}^+ \),
  
  Neumann BCs: \(-n \cdot K(\psi) \nabla(\psi + z) = g^N\), on \( \Gamma^N \times \mathbb{R}^+ \),

  where \( \partial \Omega = \Gamma^D \cup \Gamma^N \), and \( g^D \) and \( g^N \) are, respectively, Dirichlet and Neumann boundary data;

- an **initial solution**:
  
  \( \psi(x, t = 0) = \psi_0(x) \) for \( x \in \Omega \),

- the **characteristic relations** for \( \theta(\psi) \), \( K(\psi) \), and \( \eta(\psi) \) depending on soil nature: *Van Genuchten*, e.g. HYDRUS-2D, 1999, *Brooks-Corey*, Haverkamp (1977), and many other models...
Multi-layered soils are taken into consideration by assuming that \( \Omega \) be partitioned into a set of soil zones (sub-surface layers) \( \Omega_l, \ l = 1, \ldots, L \) such that

\[
\overline{\Omega} = \bigcup_{l=1,\ldots,L} \Omega_l,
\]

showing different constitutive relations:

\[
\begin{align*}
\theta(\psi)|_{\Omega_l} &= \theta^l(\psi), \\
K(\psi)|_{\Omega_l} &= K^l(\psi), \\
\eta(\psi)|_{\Omega_l} &= \eta^l(\psi).
\end{align*}
\]
The finite volume framework

Let $\mathcal{T}_h = \{ T \}$ be a suitable triangulation of $\Omega$;

Let us integrate over any control volume $T \in \mathcal{T}_h$ to obtain

$$\int_T \frac{\partial \theta(\psi)}{\partial t} \, dV - \int_T \text{div} \left[ K(\psi) \nabla (\psi + z) \right] \, dV = \int_T s \, dV;$$

then, apply the Gauss-Green divergence theorem

$$\int_T \frac{\partial \theta(\psi)}{\partial t} \, dV - \int_{\partial T} \mathbf{n} \cdot \left[ K(\psi) \nabla (\psi + z) \right] \, dS = \int_T s \, dV,$$

introduce cell averages and split face contributions to flux balance

$$\frac{d}{dt} \int_T \theta(\psi) \, dV - \sum_{e \in \partial T} |e| \int_{\partial T} \mathbf{n} \cdot \left[ K(\psi) \nabla (\psi + z) \right] \, dS = \int_T s \, dV.$$. 
Finally, let us introduce the cell-average approximation of the pressure field, e.g. \( \psi_h \equiv \{ \psi_T \} \), and the numerical flux,

\[
G_{ij}(\psi_h) = |e_{ij}| n_{ij} \cdot K_{ij} \left[ G^\cdot_{ij}(\psi_h) + \zeta \right],
\]

where \( \zeta = \nabla z \) is the versor along \( Z \),

\[
\frac{1}{K_{ij}(\psi_h)} = \frac{1}{2} \left[ \frac{1}{K(\psi_j)} + \frac{1}{K(\psi_i)} \right],
\]

and \( G^\cdot_{ij}(\psi_h) \) is the face-based gradient on the diamond \( D_{ij} \) that is given by

\[
G^\cdot_{ij}(\psi_h) = G_{ij}^{(n)}(\psi_h) n_{ij} + \frac{\psi_\beta - \psi_\alpha}{|e_{ij}|} t_{ij}.
\]
The normal component \( G_{ij}^{(n)}(\psi_h) \) of the face-based gradient \( G_{ij}^{\cdot}(\psi_h) \) is

\[
G_{ij}^{(n)}(\psi_h) = \frac{1}{|D_{ij}|} \sum_k w_{ij,k}^{(n)} \psi_k + g_{ij}^{(n)}
\]

where the weights \( \{ w_{ij,k}^{(n)} \} \) are obtained by the Least Squares reconstruction algorithm.

Rewriting

\[
\left[ G(\psi)\psi + g(\psi) \right]_i = \sum_{e_{ij} \in \partial T_i} |e_{ij}| n_{ij} \cdot K_{ij} G_{ij}^{\cdot}(\psi_h),
\]

\[
\left[ h(\psi) \right]_i = \sum_{e_{ij} \in \partial T_i} |e_{ij}| n_{ij} \cdot K_{ij} \zeta,
\]

we finally obtain the vector formulation

\[
\left. \frac{d\theta(\psi)}{dt} \right|_i - \left[ G(\psi)\psi + g(\psi) + h(\psi) \right]_i = s_i(\psi).
\]
Time discretization

- **Crucial issue**: we are advancing in time $\theta(\psi)$, but we need $\psi$ for flux balancing:

$$
\frac{d\theta(\psi)}{dt} \Bigg|_i + \left[ G(\psi)\psi + g(\psi) + h(\psi) \right]_i = s_i(\psi).
$$

- non-linear function of $\psi$
- non-linear flux term
- gravity term

- the simplest approach is to switch to a discrete form of the head-based formulation by using

$$
\frac{d\theta_i(\psi_i)}{dt} \approx \eta(\psi_i) \frac{d\psi_i}{dt}.
$$
• Using the *implicit Euler F.D. scheme* for the time derivative give us

\[
\eta(\psi^n_{i+1}) \frac{\psi^{n+1}_i - \psi^n_i}{\Delta t} - \left[ G(\psi^{n+1}_{i+1}) \psi^{n+1}_{i+1} + g(\psi^{n+1}_{i+1}) + h(\psi^{n+1}_{i+1}) \right]_i = s(\psi^{n+1}_i).
\]

• This non-linear algebraic problem can be solved by using a fixed-point iterative scheme (Picard):

\[
\left[ \text{diag}[\eta(\psi^{n+1}_{i,m})] - \Delta t G(\psi^{n+1}_{i,m}) \right] \psi^{n+1}_{i,m+1} = \\
\eta(\psi^{n+1}_{i,m}) \psi^n_i + \Delta t \left[ s_i(\psi^{n+1}_{i,m}) + g(\psi^{n+1}_{i,m}) + h(\psi^{n+1}_{i,m}) \right]
\]
• The (so-called) delta-form is obtained by solving for $\delta^{m+1,m}$

\[
\begin{align*}
\left[ \text{diag}[\eta(\psi^{n+1,m})] - \Delta tG(\psi^{n+1,m}) \right] \left( \psi^{n+1,m+1} - \psi^{n+1,m} \right) &= \\
+ \Delta t \left[ s(\psi^{n+1,m}) + g(\psi^{n+1,m}) + h(\psi^{n+1,m}) \right] \\
- \text{diag}[\eta(\psi^{n+1,m})] (\psi^{n+1,m} - \psi^n) + \Delta tG(\psi^{n+1,m}) \psi^{n+1,m}
\end{align*}
\]

\[
= \delta^{m+1,m}
\]

• Iterate on $m$ to get $\psi^{n+1,m+1}$ from $\psi^{n+1,m}$ starting from $\psi^{n+1,0} = \psi^n$ until

\[
|\delta^{m+1,m}| \leq \text{TOL}
\]

and, finally, set

\[
\psi^{n+1,m+1} = \psi^{n+1,m} + \delta^{m+1,m}
\]

This approach is easy, but shows very poor mass conservation!
What is a mass-conservative discretization?

Following M. Celia et al., W.R.R. 1990, let us define:

\[
\begin{bmatrix}
\text{total additional mass in the domain} \\
\text{total net flux into the domain}
\end{bmatrix}
= 
\begin{bmatrix}
\text{total mass in the domain at any time } t > 0 \\
\text{flux balance integrated from initial time } t = 0 \text{ to the current time } t
\end{bmatrix}
- 
\begin{bmatrix}
\text{total mass in the domain at initial time } t = 0 \\
\text{total mass in the domain at initial time } t = 0
\end{bmatrix},
\]

\[
\text{MASS BALANCE RATIO} = \frac{\text{total additional mass in the domain}}{\text{total net flux into the domain}}
\]

Clearly, it must be

\[
\text{MASS BALANCE RATIO} = 1.
\]
How does the head-based approach perform?

Let us consider, for example, the vertical column infiltration problems proposed by M. Celia et al. W.R.R. 1990:

**Case 1**

*Data:*
- column range: \([\text{bottom}] 0 \leq \zeta \leq 40 \ [\text{top}]\)
- bottom pressure: \(-61.5\) cm
- top pressure: \(-20.7\) cm
- initial pressure: \(-61.5\) cm for \(0 \leq \zeta < 40\)
- running time: from \(T = 0\) to \(T = 360\) s.

*Constitutive relationships (Haverkamp et al.):*

\[
\theta(\psi) = \frac{\alpha(\theta_s - \theta_r)}{\alpha + |\psi|^\beta} + \theta_r
\]

\[
K(\psi) = K_s \frac{A}{A + |\psi|^\gamma}
\]

\((\alpha, \beta, \gamma, \theta_s, \theta_r, K_s\text{ and }A\text{ are known parameters.})\)
Mass balance ratio for different time steps $\Delta t$

- Head-based 2D FV scheme (M. & Ferraris 2004)
- Head-based 1D FD scheme (Celia et al., 1990)
How does the head-based approach perform?

Case 2

*Data:*
- column range: \([\text{bottom}] 0 \leq \zeta \leq 100 [\text{top}]\)
- bottom pressure: \(-1000 \text{ cm}\)
- top pressure: \(-120 \text{ cm}\)
- initial pressure: \(-1000 \text{ cm for } 0 \leq \zeta < 100\)
- running time: from \(T = 0\) to \(T = 1 \text{ day}\).

*Constitutive relationships (Van Genuchten et al.)*:

\[
\theta(\psi) = \frac{\theta_s - \theta_r}{1 + (\varepsilon |\psi|^n)^m} + \theta_r
\]

\[
K(\psi) = K_s\frac{\left[1 - (\varepsilon |\psi|)^{n-1}\left[1 + (\varepsilon |\psi|^n)^{-m}\right]^{-m}\right]^2}{\left[1 + (\varepsilon |\psi|^n)^{m/2}\right]}^{m/2}
\]

\((m, n, \varepsilon, \theta_s, \theta_r, \text{ and } K_s \text{ are known parameters.})\)
Column infiltration problem: Case 2

Mass balance ratio for different time steps $\Delta t$

Head-based 2D FV scheme (M. & Ferraris, 2004)
Head-based 1D FD scheme (Celia et. al. 1990)
How can we improve this behavior?

- **Main Idea:** M. Celia et al. observed that mass conservation may be lost in the head-based formulation because of a poor approximation of the time derivative of $\theta(\psi)$:

$$\left.\frac{\partial \theta(\psi)}{\partial t}\right|_{n+1} \approx \frac{1}{|T_i|} \int_{T_i} \left.\frac{\partial \theta(\psi)}{\partial t}\right|_{i} dV \approx \frac{d}{dt} \theta(\psi_i(t))|_{n+1}$$

$$\approx \eta(\psi^n_i)\frac{\psi_i^{n+1} - \psi^n_i}{\Delta t} + O(|\delta\psi|)$$

- **Better strategy:** develop $\theta(\psi)$ to second-order in $\psi$-variation in the cell-average integral

$$\theta_i^{n+1,m+1} = \frac{1}{|T_i|} \int_{T_i} \theta(\psi_i^{n+1,m+1}) dV$$
How can we improve this behavior?

A straightforward calculation gives

\[
\theta_{i}^{n+1,m+1} = \frac{1}{|T_i|} \int_{T_i} \theta(\psi_{i}^{n+1,m+1}) \, dV
\]

\[
= \frac{1}{|T_i|} \int_{T_i} \theta\left(\psi_{i}^{n+1,m} + \left(\psi_{i}^{n+1,m+1} - \psi_{i}^{n+1,m}\right)\right) \, dV
\]

\[
= \delta_{i}^{m+1,m}
\]

\[
= \frac{1}{|T_i|} \int_{T_i} \left[ \theta(\psi_{i}^{n+1,m}) + \left| \frac{\partial \theta}{\partial \psi} \right|_{i}^{n+1,m} \left(\psi_{i}^{n+1,m+1} - \psi_{i}^{n+1,m}\right) + O(|\psi|^2) \right] \, dV
\]

\[
\approx \theta_{i}^{n+1,m} + \eta(\psi^{n+1,m}) \delta_{i}^{m+1,m} + O(|\delta \psi|^2).
\]
• Let us start again from
\[ \frac{d\theta}{dt} \bigg|_i - \left[ G(\psi)\psi + g(\psi) + h(\psi) \right]_i = s_i(\psi), \]

• and discretize directly the time derivative of \( \theta(\psi) \) to get
\[ \frac{\theta^{n+1}_i - \theta^n_i}{\Delta t} - \left[ G(\psi^{n+1})\psi^{n+1}_i + g(\psi^{n+1}) + h(\psi^{n+1}) \right]_i = s_i(\psi^{n+1}). \]

• Use Picard iterations for solving the non-linear problem
\[ \frac{\theta^{n+1,m+1}_i - \theta^n_i}{\Delta t} - \left[ G(\psi^{n+1,m})\psi^{n+1,m+1}_i + g(\psi^{n+1,m}) + h(\psi^{n+1,m}) \right]_i = s_i(\psi^{n+1,m}). \]
• Re-write Picard iterations in $\delta$-form,

\[
\theta_{i}^{n+1,m+1} - \theta_{i}^{n+1,m} = \frac{\Delta t}{\Delta t} \left[ G(\psi_{i}^{n+1,m})(\psi_{i}^{n+1,m+1} - \psi_{i}^{n+1,m}) + g(\psi_{i}^{n+1,m}) + h(\psi_{i}^{n+1,m}) \right]_{i} = \\
S_{i}(\psi_{i}^{n+1,m}) + G(\psi_{i}^{n+1,m})\psi_{i}^{n+1,m} - \frac{\theta_{i}^{n+1,m} - \theta_{i}^{n}}{\Delta t},
\]

(recall $\delta_{i}^{m+1,m} = \psi_{i}^{n+1,m+1} - \psi_{i}^{n+1,m}$).

• We will use the previously derived second-order development

\[
\theta_{i}^{n+1,m+1} \approx \theta_{i}^{n+1,m} + \eta(\psi_{i}^{n+1,m})\delta_{i}^{m+1,m} + O(|\delta\psi|^{2}).
\]
Finally, we obtain the vector expression:

\[
\begin{align*}
\text{diag} \left[ \eta(\psi^{n+1,m}) \right] - \Delta t G(\psi^{n+1,m}) \right] \delta^{m+1,m} = \\
\Delta t \left[ s(\psi^{n+1,m}) + g(\psi^{n+1,m}) + h(\psi^{n+1,m}) \right] \\
- (\theta^{n+1,m} - \theta^n) + \Delta t G(\psi^{n+1,m}) \psi^{n+1,m},
\end{align*}
\]

\[
\begin{align*}
\text{to be compared with the head-based discretization:} \\
\text{diag} \left[ \eta(\psi^{n+1,m}) \right] - \Delta t G(\psi^{n+1,m}) \right] \delta^{m+1,m} = \\
\Delta t \left[ s(\psi^{n+1,m}) + g(\psi^{n+1,m}) + h(\psi^{n+1,m}) \right] \\
- \text{diag} \left[ \eta(\psi^{n+1,m}) \right] (\psi^{n+1,m} - \psi^n) + \Delta t G(\psi^{n+1,m}) \psi^{n+1,m}.
\end{align*}
\]
Evaluation of mass conservation, Case 1

![Graph showing mass balance ratio against time-step size in sec.]

- **Mixed-based (Celia et al. 1990) 2D FV**
- **Standard Head-based 2D FV**
- **Standard Head-based 1D FD**
Evaluation of mass conservation, Case 2

- Mixed-based (Celia et.al. 1990) 2D FV
- Standard Head-based 2D FV
- Standard Head-based 1D FD
Evaluation of mass conservation, Case 1
Evaluation of mass conservation, Case 1
Evaluation of mass conservation, Case 2
Evaluation of mass conservation, Case 2

[Graph showing water content over time and depth (Z cm)]

- Water Content
- Z (cm)
- t=12h
- t=24h
- t=36h
- t=48h
The performance of the Finite Volume method for predicting unsaturated flow in multi-layered calculation is measured in two different test cases:

- horizontally layered soil (three layers), e.g. Baca, Chung, Mulla, IJNMF, 1997;
Horizontally layered soil

**Soil Data:**
- range: \(0 \text{cm} \leq \zeta \leq 25.5 \text{cm} [\text{top}]\)
- surface crust: \(25 \text{cm} \leq \zeta \leq 25.5 \text{cm}\) \(K_s = 0.0616 \text{cm h}^{-1}\)
- tilled layer: \(15 \text{cm} \leq \zeta \leq 25 \text{cm}\) \(K_s = 1.396 \text{cm h}^{-1}\)
- sub-soil: \(0 \text{cm} \leq \zeta \leq 15 \text{cm}\) \(K_s = 0.312 \text{cm h}^{-1}\)

**Constitutive relationships (Brooks-Corey et al.):**

\[
\theta(\psi) = \begin{cases} 
\theta_s & \psi \geq \psi_a, \\
\theta_s(\psi/\psi_a)^{-1/b} & \psi < \psi_a
\end{cases}
\]

\[
K(\psi) = \begin{cases} 
K_s & \psi \geq \psi_a, \\
K_s(\psi/\psi_a)^{-1/b} & \psi < \psi_a
\end{cases}
\]

(\(\theta_s\), \(\psi_a\), and \(b\) are known parameters.)
Simulation Data:

**Case 1**
- bottom pressure: $-80$ cm
- top pressure: $0$ cm
- initial pressure: $-80$ cm for $0 \leq \zeta < 25.5$ cm
- running time: from $T = 0$ to $T = 4$ hours.

**Case 2**
- bottom pressure: $-1000$ cm
- top pressure: $0$ cm
- initial pressure: $-1000$ cm for $0 \leq \zeta < 25.5$ cm
- running time: from $T = 0$ to $T = 4$ hours.
Horizontally layered soil, Case 1
Horizontally layered soil, Case 1
Horizontally layered soil, Case 2
Horizontally layered soil, Case 2

![Pressure Head (cm) vs. Z (cm) at different times](image)
Horizontally layered soil, Case 2

The graph illustrates the change in water content with depth (Z) over time (t) for different cases:

- **t=0** represents the initial water content.
- **t=1h** shows the water content after 1 hour.
- **t=2h** is the water content after 2 hours.
- **t=3h** is after 3 hours.
- **t=4h** is after 4 hours.

The water content is measured in units of Z (cm) on the x-axis, and the y-axis represents the water content range from 0.2 to 0.6.
Curvilinearly layered soil

Simulation Data

domain: $[0, L_x] \times [0, L_z]$, with $L_x = L_z 100$ cm
bottom pressure: 0 cm
top pressure: 0 cm
initial pressure: $(100 - z)$ cm for $0 \leq z < 25.5$ cm
vertical BCs: homogeneous Neumann (no flux)
running time: from $T = 0$ to $T = 2$ days.

The curvilinear layer is the curve $(x, \zeta(x))$ such that

$$\zeta(x) = \left[ \frac{1}{10} \left( 1 - \cos(\pi x / L_x) \right) + 0.45 \right] L_\zeta$$

Van Genuchten constitutive curves are used, in particular with
layer 1: $z \geq \zeta(x)$ \quad $K_s = 0.25,$
layer 2: $z \leq \zeta(x)$ \quad $K_s = 2.00.$
Curvilinearly layered soil
Curvilinearly layered soil

$T = 0$  

$T = 0.25$ day
Curvilinearly layered soil

$T = 0.5$ day

$T = 0.75$ day
Curvilinearly layered soil

$T = 1 \text{ day}$

$T = 1.25 \text{ day}$
Curvilinearly layered soil

$T = 1.5 \text{ day}$  \hspace{1cm}  $T = 1.75 \text{ day}$
Curvilinearly layered soil

\[ T = 2 \text{ day} \]
Final Remarks

- The second-order Finite Volume method based on diamond discretization of fluxes is a valid alternative to more “traditional” numerical methods like FDM, FEM, for solving flow models in partially saturated soils;

- The mixed $\theta - \psi$ formulation of the Richards equation and the mass-conservative technique by M. Celia et al. W.W.R. 1990 makes it possible the development of a very accurate time-marching method;

- The scheme is naturally adapted to multi-layered soil calculations.