

A unified modeling of underground-surface hydraulic processes with vertical-integrated approach

Carreau Manon¹ and Parisot Martin¹

¹Inria Bordeaux, Cardamom team

November 14, 2022

1 Introduction

We aim to simulate free water flow on a porous underground layer with a derived model of the Euler system and a numerical scheme already studied in [7]. This is a simplification of the water flow describe in large aquifer, here the free surface is a shallow water i.e a river or sea/ocean bank, interacting with a bathymetry bound by a substratum at the bottom and composed of a porous media or permeable materials. The substratum is an impermeable layer that can be compared with the topology used in most of the shallow water literature. The bathymetry and the substratum are both fixed in time doesn't taking into account the erosion of materials, this hypothesis is needed for the accuracy of the analysis.

This article is divided in four main sections, in §2.1 we introduce a Euler Model from which we derived the Dupuit-Forchheimer/Shallow-Water model. We recall its main proprieties. We derived the model from Euler in §2.2 with an vertical integration and then present our model in §2.2.2 with the analysis of the general proprieties of the model. This section end with the comparison with our model and common models detailed in this introduction. The numerical scheme is presented in §2.3.1 with the properties of this scheme in §2.3.2. We explain how to compute the scheme in §2.3.3. We end this article with the numerical experiments §2.4 that exhibit the proprieties of the numerical scheme.

2 The Dupuit-Forchheimer/Shallow-Water model

2.1 Framework overview

Let us start with the description of the media in which the flow is considered, namely the aquifer. A general illustration is given in Figure 1. Let $t > 0$ be the time variable, $x \in \mathbb{R}^d$ the horizontal space variable, with $d \in \{1, 2\}$ and $z \geq S(x)$ the vertical space variable. The flow is only considered over a impermeable surface $S(x) \in \mathbb{R}$, namely the substratum. Above the substratum and below a surface $B(x) \geq S(x)$, namely the surface level called bathymetry hereafter, we consider a porous media with hydrodynamic properties describe further. Note that the bathymetry is not an impermeable surface for the flow and the main motivation of this work leads in the exchanges between underground flow, i.e. below the bathymetry, with surface flow, over the bathymetry.

We assume that the flow in the whole media, i.e. $(x, z) \in \mathbb{R}^d \times [S, +\infty]$, is well described by the Euler model with a drag force. It reads

$$\begin{aligned} \partial_t s + \nabla \cdot (s\mathbf{u}) + \partial_z (sw) &= 0 \\ s(\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + w\partial_z \mathbf{u}) &= s \left(-\frac{\nabla p}{\rho} - \frac{\mathbf{u}}{\kappa} \right) \\ s(\partial_t w + \mathbf{u} \cdot \nabla w + w\partial_z w) &= s \left(-\frac{\partial_z p}{\rho} - \frac{w}{\kappa} - g \right). \end{aligned} \tag{E}$$

with the saturation $s(t, x, z) \in \mathbb{R}_+$, the horizontal velocity $\mathbf{u}(t, x, z) \in \mathbb{R}^d$ and the vertical velocity $w(t, x, z) \in \mathbb{R}$. The fluid density $\rho > 0$, the gravity acceleration $g > 0$ and the hydraulic conductivity of the media $\kappa(x, z) > 0$ are given parameters. The pressure $p(t, x, z) \in \mathbb{R}$ is an unknown taken such that the saturation does not exceed a

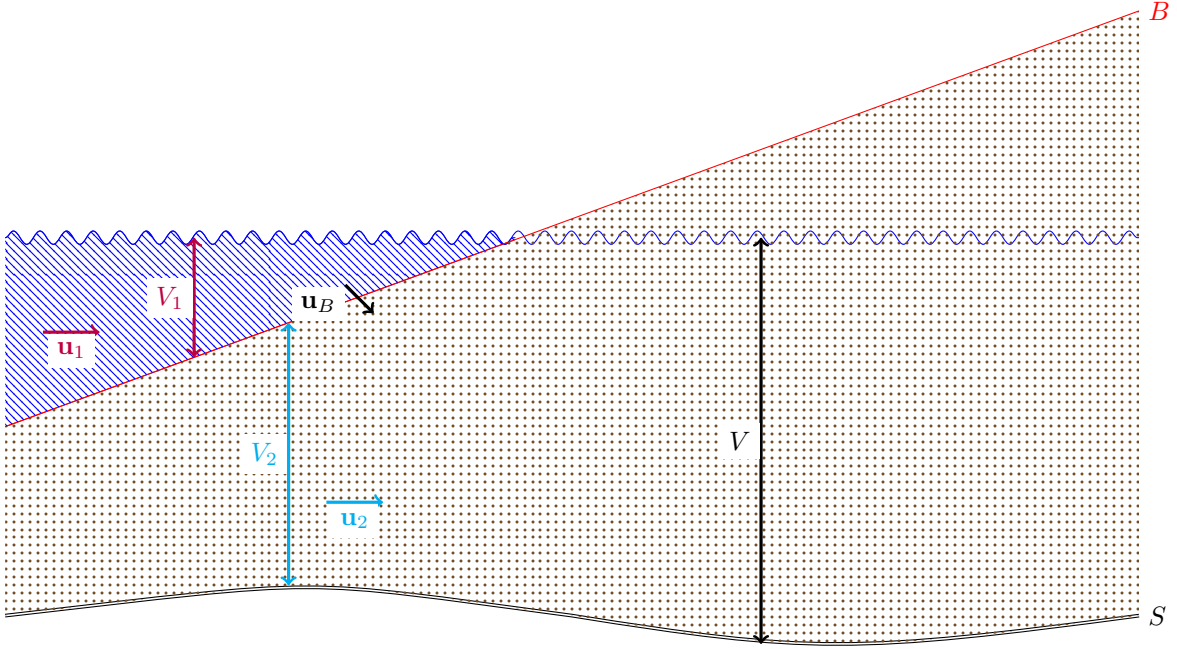


Figure 1: Illustration of problem geometry and main variables

maximum saturation, sometime called the porosity of the media, otherwise it is set to a fixed values corresponding to the atmospheric pressure neglected for simplicity i.e.

$$\min(\bar{s} - s, p) = 0 \quad (1)$$

with $0 < \bar{s}(x, z) \leq 1$ fixed, defined as the free space where the water can occupied in an infinitesimal volume. In particular, at the surface, the flow can occupied the whole space and the drag force vanished, hence $\bar{s}(x, z > S(x)) = 1$ and $\kappa(x, z > S(x)) = \infty$.

At the substratum $z = S(x)$, the non-penetration boundary condition is considered, i.e.

$$\mathbf{u}|_S \cdot \nabla S - w|_S = 0. \quad (2)$$

with the notation $\psi|_\zeta(t, x) = \psi(t, x, \zeta(x))$.

Proposition 1. *The regular enough solutions of the Euler model (E) and (1) satisfies the following energy dissipation law*

$$\partial_t E + \nabla \cdot \left(\left(\frac{p}{\rho} s + K_{\mathbf{u}} + K_w \right) \mathbf{u} \right) + \partial_z \left(\left(\frac{p}{\rho} s + K_{\mathbf{u}} + K_w \right) w \right) = -\frac{s}{\kappa} (|\mathbf{u}|^2 + |w|^2)$$

with the kinetic energy $K_{\mathbf{u}}$ and K_w for the horizontal and vertical velocities respectively defined as

$$K_{\mathbf{u}} = s \frac{|\mathbf{u}|^2}{2} \text{ and } K_w = s \frac{|w|^2}{2}$$

and the total energy $E = \mathcal{E}_E + K_{\mathbf{u}} + K_w$, \mathcal{E}_E the potential energy such as $\mathcal{E}_E(s, z) := gzs$.

Proof. We start by multiplying the mass equation by the potential gz to find the potential energy equation.

$$\partial_t \mathcal{E} + \nabla \cdot (gzs\mathbf{u}) + \partial_z (gzs w) = swg \quad (3)$$

Then we write the two momentum equations under their conservative form, using the mass conservation from the (E) system

$$\partial_t (s\mathbf{u}) + \nabla \cdot (s\mathbf{u} \otimes \mathbf{u}) + \partial_z (sw\mathbf{u}) = s \left(-\frac{\nabla p}{\rho} - \frac{\mathbf{u}}{\kappa} \right)$$

$$\partial_t (sw) + \nabla \cdot (s\mathbf{u}w) + \partial_z (s|w|^2) = s \left(-\frac{\partial_z p}{\rho} - \frac{w}{\kappa} - g \right).$$

Adding the conservative and non-conservative form of the horizontal (resp. vertical) momentum equation and multiplying by $\frac{w}{2}$ (resp. $\frac{w}{2}$) we obtain

$$\partial_t \left(s \frac{|\mathbf{u}|^2}{2} \right) + \nabla \cdot \left(s \frac{|\mathbf{u}|^2}{2} \otimes \mathbf{u} \right) + \frac{1}{\rho} \nabla \cdot (s\mathbf{u}p) + \partial_z \left(s \frac{|\mathbf{u}|^2}{2} w \right) = \frac{p}{\rho} \nabla \cdot (s\mathbf{u}) - s \frac{|\mathbf{u}|^2}{\kappa} \quad (4)$$

$$\partial_t \left(s \frac{|w|^2}{2} \right) + \nabla \cdot \left(s \frac{|w|^2}{2} \mathbf{u} \right) + \partial_z \left(s \frac{|w|^2}{2} w \right) + \frac{1}{\rho} \partial_z (swp) = \frac{p}{\rho} \partial_z (sw) - s \frac{|w|^2}{\kappa} - swg. \quad (5)$$

Summing (3), (4) and (5), it remains to estimate the term multiplied by the pressure p . Using the mass equation multiplied we get,w

$$\begin{aligned} p(\nabla \cdot (s\mathbf{u}) + \partial_z (sw)) &= p(\partial_t (\bar{s} - s)) \\ &= \partial_t (p(\bar{s} - s)) - (\bar{s} - s) \partial_t p \end{aligned}$$

Using the complementarity condition (1),

$$0 = \partial_t (p(\bar{s} - s)) = p \partial_t (\bar{s} - s) + (\bar{s} - s) \partial_t p.$$

either $p = 0$ or $(\bar{s} - s) = 0$, so we can conclude $p \partial_t (\bar{s} - s) = 0$ and our result. \square

2.2 Derivation of a unified model for an aquifer

The goal of this section is to derive a reduced model that takes recover classical models in each particular areas, mainly the shallow-water model [2] in the free surface subdomain and the Dupuit-Forchheimer model [3] in the porous media subdomain. This derivation is realized in two steps. First using a vertical integration, we recover a bilayer model which split the two subdomains. Then assuming the hydraulic conductivity small enough in the porous media, we finally obtain a unified model.

2.2.1 Vertical integration

The two classical models mentioned, i.e. the shallow-water model and the Dupuit-Forchheimer model, basically assume the same hypothesis that we recall here.

Hypothesis 1 (mono-valuated free surface). *Assume that there exists an upper surface $\eta(t, x) \geq S(x)$ such that the media below the upper surface is saturated and the media above the upper surface is empty of water, i.e.*

$$s(t, x, z) = \begin{cases} \bar{s}(x, z) & \text{if } S(x) \leq z \leq \eta(t, x), \\ 0 & \text{else.} \end{cases}$$

In the framework of the underground flows, this upper surface is called hydraulic table where as it is called free surface in the framework of free surface flows. Remark that it is an unknown of the study. This assumption imply that the unsaturated zones are not taken into account in this study. It will be treated in a further work.

It is convenient to introduce the volume of water in each sub domain, i.e.

$$V_f(t, x) := \int_{B(x)}^{\infty} s(t, x, z) \, dz \quad \text{and} \quad V_g(t, x) := \int_{S(x)}^{B(x)} s(t, x, z) \, dz$$

and the free volume under the bathymetry as

$$\bar{V}_g(x) = \int_{S(x)}^{B(x)} \bar{s}(x, z) \, dz. \quad (6)$$

Now remark the following property.

Proposition 2. *Assume that Hypothesis 1 is satisfied. Then the following complementarity condition holds*

$$\min(\bar{V}_g - V_g, V_f) = 0.$$

Proof. Since the saturation is bounded (1), it follows that $V_g(t, x) \leq \bar{V}_g(x)$ for any $t \geq 0$ and $x \in \mathbb{R}^d$.

In addition, for a given t and x if $V_g(t, x) \neq \bar{V}_g(x)$, we have $s(t, x, z = B(x)) = 0$ accordingly to Hypothesis 1. We conclude that $s(t, x, z \geq B(x)) = 0$, hence $V_f(t, x) = 0$. \square

The complementary condition Proposition 2 imply that there is a unique unknown $V = V_f + V_g$. We can now introduce two new functions V_1 and V_2 that respect the complementarity condition (2) and such as $V = V_1 + V_2$.

$$V_2(V, x) = \min(\bar{V}_2(x), V) \quad \text{and} \quad V_1(V, x) = V - V_2(V, x). \quad (7)$$

We now address the question to write the reduced model of the evolution of the total volume of water V . As it will be show further, under classical assumptions and following the strategy of layer wise vertical integration, see [1], the 3D Euler model (E) can be reduced to the following bi-layer model

$$\begin{aligned} \partial_t V_1 + \nabla \cdot (V_1 \bar{\mathbf{u}}_1) &= -G \\ \partial_t V_2 + \nabla \cdot (V_2 \bar{\mathbf{u}}_2) &= G \\ \partial_t (V_1 \bar{\mathbf{u}}_1) + \nabla \cdot (V_1 \bar{\mathbf{u}}_1 \otimes \bar{\mathbf{u}}_1) &= -gV_1 \nabla (h + S) - \frac{V_1 \bar{\mathbf{u}}_1}{\bar{\kappa}_1} - \mathbf{u}_B G \\ \partial_t (V_2 \bar{\mathbf{u}}_2) + \nabla \cdot (V_2 \bar{\mathbf{u}}_2 \otimes \bar{\mathbf{u}}_2) &= -gV_2 \nabla (h + S) - \frac{V_2 \bar{\mathbf{u}}_2}{\bar{\kappa}_2} + \mathbf{u}_B G \end{aligned} \quad (SW_2)$$

where the unknowns are the depth of the water column $h(t, x) := \eta(t, x) - S(x) \in \mathbb{R}_+$, the mass exchanged $G(t, x) \in \mathbb{R}$ which ensure that the complementarity condition, Proposition 2, is satisfied, and the averaged horizontal velocities in each sub domain $\bar{\mathbf{u}}_1(t, x) \in \mathbb{R}^d$ and $\bar{\mathbf{u}}_2(t, x) \in \mathbb{R}^d$.

Note that $\bar{\mathbf{u}}_1$ (resp. $\bar{\mathbf{u}}_2$) is not defined when $V_1 = 0$ (resp. $V_2 = 0$). We set $\bar{\mathbf{u}}_1 = 0$ (resp. $\bar{\mathbf{u}}_2 = 0$) when $V_1 = 0$ (resp. $V_2 = 0$). The total volume of water V , that give V_1 and V_2 see (7), can be recovered from the water depth h knowing the maximum of saturation. With a small abus of notation used for readability, the total height of water $h(V, x)$ is choose such that

$$\int_{S(x)}^{S(x)+h(V,x)} \bar{s}(x, z) \, dz = V. \quad (8)$$

The parameters $\bar{\kappa}_1(x) > 0$ and $\bar{\kappa}_2(x) > 0$ are respectively the effective hydraulic conductivities in each subdomain. The velocity $\mathbf{u}_B(t, x) \in \mathbb{R}^d$ is the horizontal velocity at the interface $z = B$. As for other layerwise model, this velocity have to be given from the unknowns. In this work, we use the linear reconstruction

$$\mathbf{u}_B = \frac{\bar{\mathbf{u}}_1 + \bar{\mathbf{u}}_2}{2} + \lambda \text{sign}(G) \frac{(\bar{\mathbf{u}}_1 - \bar{\mathbf{u}}_2)}{2} \quad (9)$$

with $\lambda \in \mathbb{R}$ is a parameter set to ensure the energy stability see Proposition 4.

Let us first details the assumptions required to obtain the reduced model (7), (SW₂) and (8) from (E). For this we introduce the small parameter $\varepsilon > 0$ which represents the error of modeling.

Hypothesis 2 (Uniform hydraulic conductivity). *Assume that the hydraulic conductivity $\kappa(x, z)$ is close enough from the effective hydraulic conductivities under the bathymetry and over the bathymetry, i.e.*

$$\kappa(x, z) = \begin{cases} \bar{\kappa}_1(x) + O(\varepsilon^2) & \text{if } S(x) \leq z \leq B(x), \\ \bar{\kappa}_2(x) + O(\varepsilon^2) & \text{if } B(x) \leq z. \end{cases}$$

Hypothesis 3 (Hydrostatic pressure). *Assume that the pressure is mainly governed by the gravity, i.e.*

$$p = \rho g (\eta - z) + O(\varepsilon^2).$$

This assumption basically imply that the vertical velocity is negligible. It is known as the Dupuit-Forchheimer assumption in the framework of underground flows and the hydrostatic assumption for free surface flows. Eventually, we define the vertical-averaged horizontal velocity in each subdomain as follow

$$\bar{\mathbf{u}}_1(t, x) = \frac{1}{V_1} \int_B^\infty \mathbf{s} \mathbf{u} \, dz \quad \text{and} \quad \bar{\mathbf{u}}_2(t, x) = \frac{1}{V_2} \int_S^B \mathbf{s} \mathbf{u} \, dz.$$

Hypothesis 4 (Uniform horizontal velocity). *Assume that the horizontal velocity is close enough from its vertical-averaged in each subdomains under the bathymetry and over the bathymetry, i.e.*

$$s\mathbf{u}(t, x, z) = \begin{cases} s\bar{\mathbf{u}}_1(t, x) + O(\varepsilon) & \text{if } S(x) \leq z \leq B(x), \\ s\bar{\mathbf{u}}_2(t, x) + O(\varepsilon) & \text{if } B(x) \leq z. \end{cases}$$

Remark that this assumption is not required for underground flow using the traditional derivation but it is recovered at the end. It is required for the derivation of the shallow-water model but can be also justified in specific regimes [4].

Proposition 3. *Assume that Hypothesis 1, Hypothesis 2, Hypothesis 3 and Hypothesis 4 are fulfilled. Then the bi-layer model (7), (SW₂) and (8) is an approximation in $O(\varepsilon^2)$ of the Euler model (E).*

Proof. Let us first consider the mass equation of the Euler system (E). By integration between the substratum and the bathymetry and using a Leibniz rule, we get

$$0 = \int_S^B (\partial_t s + \nabla \cdot (s\mathbf{u}) + \partial_z (sw)) dz = \partial_t V_2 + \nabla \cdot (V_2 \mathbf{u}_2) - s|_B (\mathbf{u}|_B \cdot \nabla B - w|_B) + s|_S (\mathbf{u}|_S \cdot \nabla S - w|_S).$$

The last term vanish thanks to the no-penetration condition (2). We define the vertical mass flux by $G := s|_B (\mathbf{u}|_B \cdot \nabla B - w|_B)$. Note that this flux should vanish when $s|_B = 0$, i.e. if $\eta < B$. We get the mass balance of the underground layer (second equation of (SW₂)). Similarly by integration the mass equation of the Euler system (E) between the bathymetry to infinity, we get the mass balance of the surface layer (first equation of (SW₂)). Now we focus on the momentum equation of the Euler system (E) written under its conservative form using the mass equation. By integration between the substratum and the bathymetry, we get

$$\begin{aligned} 0 &= \int_S^B \left(\partial_t (s\mathbf{u}) + \nabla \cdot (s\mathbf{u} \otimes \mathbf{u}) + \partial_z (s\mathbf{u}w) - \frac{s\nabla p}{\rho} - \frac{s\mathbf{u}}{\kappa} \right) dz \\ &= \partial_t (V_2 \bar{\mathbf{u}}_2) + \nabla \cdot \left(\int_S^B s\mathbf{u} \otimes \mathbf{u} dz \right) - \mathbf{u}|_B G - \int_S^B \frac{s\nabla p}{\rho} dz - \int_S^B \frac{s\mathbf{u}}{\kappa} dz. \end{aligned}$$

Using Hypothesis 4 and setting $\varepsilon \tilde{\mathbf{u}}_2 = \mathbf{u} - \bar{\mathbf{u}}_2$, we have

$$\int_S^B s\mathbf{u} \otimes \mathbf{u} dz = \int_S^B s(\bar{\mathbf{u}}_2 + \varepsilon \tilde{\mathbf{u}}_2) \otimes (\bar{\mathbf{u}}_2 + \varepsilon \tilde{\mathbf{u}}_2) dz = V_2 \bar{\mathbf{u}}_2 \otimes \bar{\mathbf{u}}_2 + \varepsilon \left(\bar{\mathbf{u}}_2 \otimes \int_S^B s\tilde{\mathbf{u}}_2 dz + \int_S^B s\tilde{\mathbf{u}}_2 dz \otimes \bar{\mathbf{u}}_2 \right) + O(\varepsilon^2).$$

The term in $O(\varepsilon)$ vanishes since by construction, i.e. $\int_S^B s\mathbf{u} dz = \int_S^B s\bar{\mathbf{u}}_2 dz$ hence $\int_S^B s\tilde{\mathbf{u}}_2 dz = 0$.

Then using Hypothesis 3, the horizontal gradient of the pressure does not depend on the vertical coordinate. More precisely, we have

$$\int_S^B \frac{s\nabla p}{\rho} dz = \int_S^B s dz g\nabla\eta + O(\varepsilon^2) = gV_2\nabla\eta + O(\varepsilon^2).$$

Finally, using Hypothesis 2 and setting $\varepsilon \tilde{\kappa}_2 = \kappa - \bar{\kappa}_2$, the last term reads

$$\int_S^B \frac{s\mathbf{u}}{\kappa} dz = \frac{\int_S^B s\mathbf{u} dz}{\bar{\kappa}_2} + O(\varepsilon^2) = \frac{V_2 \mathbf{u}_2}{\bar{\kappa}_2} + O(\varepsilon^2).$$

We proceed similarly for the momentum balance of the surface layer. □

Remark 1. *As seen in the demonstration of Proposition 3, the mass exchange G should vanish (locally) as long as $\eta < B$. It is actually the case since the total volume V have to reach \bar{V}_2 before the volume in the surface layer increase thanks to (7). More precisely, considering the time derivative of the complementarity equation Proposition 2, we conclude that when $V < \bar{V}_2$, we have $\partial_t V_1 = 0$, hence $G = 0$ thanks to the mass balance in the surface layer.*

Let us now define the potential of the conservative forces by $\phi := g(h + S)$. Before going further in the reduction of the model, let us highlight the following energy balance of the bi-layer model.

Proposition 4. For any $\lambda \geq 0$, the regular enough solutions of the bilayer model (7), (SW₂) and (8) with the closure (9) satisfies the following energy dissipation law

$$\partial_t \left(\mathcal{E}_{SW_2} + \sum_{i=1}^2 V_i \frac{|\bar{\mathbf{u}}_i|^2}{2} \right) + \nabla \cdot \left(\sum_{i=1}^2 \left(g(h+S) + \frac{|\bar{\mathbf{u}}_i|^2}{2} \right) V_i \bar{\mathbf{u}}_i \right) \leq - \sum_{i=1}^2 \frac{V_i}{\bar{\kappa}_i} \|\bar{\mathbf{u}}_i\|^2.$$

with $\mathcal{E}_{SW_2}(V, x) = g \int_{S(x)}^{S(x)+h(V,x)} z \bar{s}(x, z) \, dz$

Proof. This proof is an adaptation of the result given in [1, Proposition 4.2]. Let us start by multiplying the mass equations by the potential $\phi = g(h+S)$, we get the potential energy balance, i.e.

$$0 = \phi (\partial_t V + \nabla \cdot (V_1 \bar{\mathbf{u}}_1 + V_2 \bar{\mathbf{u}}_2)) = \partial_t \mathcal{E}_{SW_2} + \nabla \cdot (\phi (V_1 \bar{\mathbf{u}}_1 + V_2 \bar{\mathbf{u}}_2)) - V_1 \bar{\mathbf{u}}_1 \cdot \nabla \phi - V_2 \bar{\mathbf{u}}_2 \cdot \nabla \phi, \quad (10)$$

with $\partial_t \mathcal{E}_{SW_2} = \phi \partial_t V$.

Now let us start by writing the momentum equations under the non conservative form, i.e. for the surface layer

$$V_1 (\partial_t \bar{\mathbf{u}}_1 + \bar{\mathbf{u}}_1 \cdot \nabla \bar{\mathbf{u}}_1) = -V_1 \nabla \phi - \frac{V_1 \bar{\mathbf{u}}_1}{\bar{\kappa}_1} - (\mathbf{u}_B - \bar{\mathbf{u}}_1) G$$

Summing the conservative and the non-conservative form and multiplying by the corresponding horizontal velocity, we get the kinetic energy balance, i.e.

$$\partial_t \left(V_1 \frac{|\bar{\mathbf{u}}_1|^2}{2} \right) + \nabla \cdot \left(\frac{|\bar{\mathbf{u}}_1|^2}{2} V_1 \bar{\mathbf{u}}_1 \right) = -V_1 \bar{\mathbf{u}}_1 \cdot \nabla \phi - \frac{V_1}{\bar{\kappa}_1} \|\bar{\mathbf{u}}_1\|^2 - \bar{\mathbf{u}}_1 \cdot \left(\mathbf{u}_B - \frac{\bar{\mathbf{u}}_1}{2} \right) G. \quad (11)$$

Proceeding similarly with the underground layer we get

$$\partial_t \left(V_2 \frac{|\bar{\mathbf{u}}_2|^2}{2} \right) + \nabla \cdot \left(\frac{|\bar{\mathbf{u}}_2|^2}{2} V_2 \bar{\mathbf{u}}_2 \right) = -V_2 \bar{\mathbf{u}}_2 \cdot \nabla \phi - \frac{V_2}{\bar{\kappa}_2} \|\bar{\mathbf{u}}_2\|^2 + \bar{\mathbf{u}}_2 \cdot \left(\mathbf{u}_B - \frac{\bar{\mathbf{u}}_2}{2} \right) G. \quad (12)$$

In order to get the result by summing (10), (11) and (12), it remains to estimate the contribution of the vertical exchanges, i.e. the last term of (11) and (12). Using the closure (9), we write

$$\left(\bar{\mathbf{u}}_2 \cdot \left(\mathbf{u}_B - \frac{\bar{\mathbf{u}}_2}{2} \right) - \bar{\mathbf{u}}_1 \cdot \left(\mathbf{u}_B - \frac{\bar{\mathbf{u}}_1}{2} \right) \right) G = -\lambda \frac{\|\bar{\mathbf{u}}_2 - \bar{\mathbf{u}}_1\|^2}{2} |G|$$

which is negative under the assumption of the proposition. \square

2.2.2 Low-conductivity asymptotic limit

In order to go further in the model reduction, one can look at particular parameters setting relevant to geophysical flows. More precisely, we assume that the media is stratified such that the porosity is constant by layer, i.e.

$$\bar{s}(x, z) = \begin{cases} \bar{s}_1(x) & \text{if } B \leq z, \\ \bar{s}_2(x) & \text{if } S \leq z < B. \end{cases} \quad (13)$$

with $(\bar{s}_1, \bar{s}_2) \in]0, 1]^2$. The unknowns can be defined from the volume of water $V(t, x)$

$$h(V, x) = \frac{V - \min(\bar{V}_2, V)}{\bar{s}_1} + \frac{\min(\bar{V}_2, V)}{\bar{s}_2} \quad \text{and} \quad \mathcal{E}(V, x) = g \left(V_1 \left(\frac{V_1}{2 \bar{s}_1} + B \right) + V_2 \left(\frac{V_2}{2 \bar{s}_2} + S \right) \right) \quad (14)$$

with $\bar{V}_2(x) = \bar{s}_2(B(x) - S(x))$, V_1 and V_2 defined by (7). $\mathbb{1}_{V \leq \bar{V}_2}$ (resp. $\mathbb{1}_{V > \bar{V}_2}$) are the indicator function such as $\mathbb{1}_{V \leq \bar{V}_2} = 1$ if $V \leq \bar{V}_2$ and 0 if not (resp. $\mathbb{1}_{V > \bar{V}_2} = 1$ if $V > \bar{V}_2$ and 0 if not). In addition, we assume that the hydraulic conductivity is low in the porous media and infinite for the surface layer, i.e.

$$\bar{\kappa}_1(x) = 0 \left(\frac{1}{K_2} \right) \quad \text{and} \quad \bar{\kappa}_2(x) = O(K_2) \quad (15)$$

with a small parameter K_2 .

With an abus of notation used to highlight the link with the bi-layer model, the shallow-water/Dupuit-Forchheimer model reads

$$\begin{aligned} \partial_t V_1 + \nabla \cdot (V_1 \bar{\mathbf{u}}_1) &= -G \\ \partial_t V_2 - \nabla \cdot (g \bar{\kappa}_2 V_2 \nabla (h + S)) &= G \\ \partial_t (V_1 \bar{\mathbf{u}}_1) + \nabla \cdot (V_1 \bar{\mathbf{u}}_1 \otimes \bar{\mathbf{u}}_1) &= -V_1 \nabla \phi - \mathbf{u}_B G \end{aligned} \quad (SW/DF)$$

with (7), (8) and the closure of the velocity at the interface reads

$$\mathbf{u}_B = \frac{1 + \lambda \operatorname{sign}(G)}{2} \bar{\mathbf{u}}_1. \quad (16)$$

Proposition 5. *Assume the parameters satisfy the scaling (13) and (15).*

Then the model (7), (8), (SW/DF) and (16) is an approximation in $O(K_2^2)$ of the bi-layer model (7), (8), (SW₂) and (9).

Proof. Lets V_2 and $\bar{\mathbf{u}}_2$ be expand as Hilbert series as follow

$$V_2 = \sum_{i=0}^{+\infty} K_2^i V_2^{(i)}, \quad \bar{\mathbf{u}}_2 = \sum_{i=0}^{+\infty} K_2^i \bar{\mathbf{u}}_2^{(i)} \quad \text{and} \quad G = \sum_{i=0}^{+\infty} K_2^i G^{(i)}.$$

Injecting the expansion in the momentum equation of the underground layer of (SW₂), the term of main order formally reads

$$-V_2^{(0)} \bar{\mathbf{u}}_2^{(0)} = 0$$

and we conclude that $\bar{\mathbf{u}}_2^{(0)} = 0$ since by definition $\bar{\mathbf{u}}_2 = 0$ when $V_2 = 0$. Similarly, the second order term formally reads

$$0 = -g V_2^{(0)} \nabla (h^{(0)} + S) - \frac{V_2^{(0)} \bar{\mathbf{u}}_2^{(1)}}{\bar{\kappa}_2^{(1)}} + \mathbf{u}_B^{(0)} G^{(0)}, \quad (17)$$

with $\bar{\kappa}_2^{(1)} = \frac{\bar{\kappa}_2}{K_2}$. Now we focus on the estimation of the term $G^{(0)}$. Injecting the expansion in the mass equation of the underground layer of (SW₂), the term of main order formally reads

$$\partial_t V_2^{(0)} = G^{(0)}.$$

For (t, x) such that $V(t, x) < \bar{V}_2(x)$, we have $G(t, x) = 0$ (see Remark 1).

For (t, x) such that $V(t, x) > \bar{V}_2(x)$, using the same arguments, i.e. considering the time derivative of the complementarity equation Proposition 2, we have that $\partial_t V_2 = 0$, hence $G^{(0)} = 0$.

For (t, x) such that $V(t, x) = \bar{V}_2(x)$, the time derivative complementarity equation is less easy to use. It only say that one of the two derivatives vanish, i.e. $\partial_t V_1 \partial_t V_2 = 0$ (for regular enough functions). If $\partial_t V_2 = 0$ then $G^{(0)} = 0$. Otherwise $\partial_t V_1 = 0$ and $V_1 = 0$, hence the mass equation of the surface layer in (SW₂) leads to $G = 0$.

Anyway we conclude that $G^{(0)} = 0$. Back to (17), we conclude that

$$V_2 \bar{\mathbf{u}}_2 = -g \bar{\kappa}_2 V_2 \nabla (h + S) + O(K_2^2)$$

That give the mass equation of the underground layer in (SW/DF). In addition, since we have $G^{(0)} = 0$, it means that the vertical flux $G = O(K_2)$. Hence, we can neglect the terms in $O(K_2)$ in (9), in particular the velocity $\bar{\mathbf{u}}_2$, and keep a model in $O(K_2^2)$, that give the closure (16). \square

One step further in the modeling reduction is to consider a particular scaling in the surface layer. In the context of the simulation of underground and surface flow in an aquifer, one can consider that the flow at the surface is completely free, i.e.

$$\bar{\kappa}_1 \rightarrow \infty \quad \text{and} \quad \bar{s}(x, z > B(x)) = 1.$$

This scaling does not really simplified the model except that the drag term in the surface layer negligible. However, this trivial simplification of (SW/DF) seems relevant in practice.

Proposition 6. For any $\lambda \geq 0$ the regular enough solutions of the shallow-water/Dupuit-Forchheimer (7), (SW/DF) and (8) with the closure (16) satisfies the following energy dissipation law

$$\partial_t \left(\mathcal{E} + V_1 \frac{|\bar{\mathbf{u}}_1|^2}{2} \right) + \nabla \cdot \left(\phi (V_1 \bar{\mathbf{u}}_1 - \bar{\kappa}_2 V_2 \nabla \phi) + V_1 \bar{\mathbf{u}}_1 \frac{|\bar{\mathbf{u}}_1|^2}{2} \right) \leq -\kappa_2 V_2 |\nabla \phi|^2$$

with \mathcal{E} defined as in (14).

Proof. This proof is analogous to the one in the Proposition 4. We remind about the potential $\phi = g(h + S)$, and add the two mass equations of (SW/DF) before multiplying the result by ϕ to find the potential energy balance

$$0 = \phi (\partial_t V + \nabla \cdot (V_1 \bar{\mathbf{u}}_1 - \bar{\kappa}_2 V_2 \nabla \phi)) = \partial_t \mathcal{E} + \nabla \cdot (\phi (V_1 \bar{\mathbf{u}}_1 - \bar{\kappa}_2 V_2 \nabla \phi)) - V_1 \bar{\mathbf{u}}_1 \cdot \nabla \phi + \bar{\kappa}_2 V_2 \cdot |\nabla \phi|^2. \quad (18)$$

The kinetic energy balance equation is exactly the same than in the Bi-layer energy balance proof (11).

$$\partial_t \left(V_1 \frac{|\bar{\mathbf{u}}_1|^2}{2} \right) + \nabla \cdot \left(\frac{|\bar{\mathbf{u}}_1|^2}{2} V_1 \bar{\mathbf{u}}_1 \right) = -V_1 \bar{\mathbf{u}}_1 \cdot \nabla \phi - \bar{\mathbf{u}}_1 \cdot \left(\mathbf{u}_B - \frac{\bar{\mathbf{u}}_1}{2} \right) G.$$

It remains to calculate the last term of the right hand side of the kinetic energy balance equation with (16)

$$-\bar{\mathbf{u}}_1 \cdot \left(\mathbf{u}_B - \frac{\bar{\mathbf{u}}_1}{2} \right) G = -\bar{\mathbf{u}}_1 \cdot \left(\frac{1 + \lambda \operatorname{sgn}(G)}{2} \bar{\mathbf{u}}_1 - \frac{\bar{\mathbf{u}}_1}{2} \right) G = -\lambda \frac{|\bar{\mathbf{u}}_1|^2}{2} |G|,$$

and we can conclude that these term is negative under the assumption of this proposition. Finally the result is found by summing (18) and (11). \square

2.2.3 Link with classical models in geophysics

This section present situations where the (SW/DF) model is similar in comparison to other classical models. The (SW/DF) model is used in situations with two layers, one underground and a free surface. For extreme situation with one of layer nonexistent, the model is analogous to the one layer Shallow Water model

$$\begin{aligned} \partial_t h + \nabla \cdot (h \mathbf{u}) &= 0 \\ \partial_t (h \mathbf{u}) + \nabla \cdot \left(h \mathbf{u} \times \mathbf{u} + g \frac{h^2}{2} \right) &= -gh \nabla S \end{aligned} \quad (SW)$$

with h the height of the water, \mathbf{u} the velocity vector and ϕ the potential or the nonlinear Dupuit-Forchheimer equation

$$\mu \partial_t h - g \nabla \cdot (\kappa h \nabla (h + S)) = 0 \quad (DF)$$

with h the water height, κ the hydraulic conductivity, μ the porosity and S the substratum.

Theorem 1. If $B = S$ and $\bar{s}_1 \in \mathbb{R}$ then the solutions of (SW/DF) and (SW) are equivalent.

Proof. $B = S$ implies there is no underground layer, $V_2 = 0$, $V_1 = V$, and $h = \frac{V}{\bar{s}_1}$. It means after dividing by the saturation \bar{s}_1 , (SW/DF) can be rewritten

$$\begin{aligned} \partial_t h + \nabla \cdot (h \bar{\mathbf{u}}_1) &= 0 \\ \partial_t (h \bar{\mathbf{u}}_1) + \nabla \cdot (h \bar{\mathbf{u}}_1 \times \bar{\mathbf{u}}_1) &= -gh \nabla (h + S) \end{aligned}$$

After separating the term of the right hand side of the momentum equation, we find the model (SW). \square

Theorem 2. If $\kappa_2 = 0$, $\bar{s}_1 \in \mathbb{R}$ and $V > \bar{V}_2$ then the solutions of (SW/DF) and (SW) are equivalent.

Proof. $\kappa_2 = 0$ implies that $\partial_t V_2 = G$, and $V > \bar{V}_2$ means $\partial_t \bar{V}_2 = 0$ and $G = 0$. It implies there is no mass exchange between the two layers.

$$\begin{aligned} \partial_t V_1 + \nabla \cdot (V_1 \bar{\mathbf{u}}_1) &= 0 \\ \partial_t (V_1 \bar{\mathbf{u}}_1) + \nabla \cdot (V_1 \bar{\mathbf{u}}_1 \times \bar{\mathbf{u}}_1) &= -g V_1 \nabla (h + S) \end{aligned}$$

In this situation $h = \frac{V - \bar{V}_2}{\bar{s}_1} + \frac{\bar{V}_2}{\bar{s}_2}$, it means $h + S = \frac{V - \bar{V}_2}{\bar{s}_1} + B$. Dividing by \bar{s}_1 supposed to be constant, we got

$$\begin{aligned} \partial_t \left(\frac{V - \bar{V}_2}{\bar{s}_1} \right) + \nabla \cdot \left(\frac{V - \bar{V}_2}{\bar{s}_1} \bar{\mathbf{u}}_1 \right) &= 0 \\ \partial_t \left(\frac{V - \bar{V}_2}{\bar{s}_1} \bar{\mathbf{u}}_1 \right) + \nabla \cdot \left(\frac{V - \bar{V}_2}{\bar{s}_1} \bar{\mathbf{u}}_1 \times \bar{\mathbf{u}}_1 \right) &= -g \frac{V - \bar{V}_2}{\bar{s}_1} \nabla \left(\frac{V - \bar{V}_2}{\bar{s}_1} + B \right) \end{aligned}$$

Noting $h_1 = \frac{V - \bar{V}_2}{\bar{s}_1}$, we find the model (*SW*) with $h = h_1$ and the topography $S = B$. □

Remark 2. *The saturation \bar{s}_1 is constant for this two theorems, it's a restrictive hypothesis but a physical one. Here, \bar{s}_1 is the saturation of the free surface layer, in practice, it means that $\bar{s}_1 = 1$ for the whole layer.*

Theorem 3. *If $V < \bar{V}_2$ then there is only the underground layer, and the solutions of (*SW/DF*) and (*DF*) are equivalent.*

Proof. For $V < \bar{V}_2$, $V_1 = 0$, $V = V_2$ and the mass exchange term $G = 0$, (*SW/DF*) can be rewrite

$$\partial_t V - \nabla \cdot (g \bar{\kappa}_2 V \nabla (h + S)) = 0$$

In this situation, $h = \frac{V}{\bar{s}_2}$ and the equation become

$$\bar{s}_2 \partial_t h - \nabla \cdot (g \bar{\kappa}_2 \bar{s}_2 h \nabla (h + S)) = 0$$

Choosing $\mu = \bar{s}_2$ and $\kappa = \bar{\kappa}_2 \bar{s}_2$ we find the model (*DF*). □

2.3 Numerical resolution

This section present the numerical scheme used to obtain an approximation of the solution of the (*SW/DF*) model. This numerical scheme is an adaptation of the Centered-Potential Regularization method (CPR) describe in [7].

2.3.1 Numerical Scheme

For any dimension $d \in \mathbb{N} \setminus \{0\}$, the mesh \mathbb{T} is composed by polygonal called control volumes, we note m_k the surface area of the k^{th} control volume and \mathbb{F}_k the set of its faces. For each face $f \in \mathbb{F}_k$, we note by m_f the length of the face, $\mu_f^k = \frac{m_f}{\sum_{j \in \mathbb{F}_k} m_j}$ the normalized length, k_f the control volume such that $f = k \cap k_f$, $N_k^{k_f}$ the unit normal to the face f outward to the control volume k . The compactness of the k^{th} control volume is noted $\delta_k = \frac{m_k}{\sum_{f \in \mathbb{F}_k} m_f}$.

For the following scheme, we used the following reconstruction at the face, $(a)_f = \frac{a_k + a_{k_f}}{2}$ and $[a]_k^{k_f} = \frac{a_{k_f} - a_k}{2}$. The numerical unknowns of the CPR scheme is the potential of the conservative forces ϕ_k^n and the horizontal velocity $\bar{\mathbf{u}}_k^n$, the subscript k means an approximation in the cell k of ϕ and $\bar{\mathbf{u}}$ at time $t^n = t^{n-1} + \delta_t^n$. With the same logic, we denote x_k an approximation in the cell k of the positional coordinate x . The time step δ_t^n is adapted such that a CFL condition ensuring the stability of the scheme is satisfied, see §2.3.2. This strategy allows the scheme to be robust to the regime changes, see [5, 6]. The total water volume are recover from the potential by

$$V(\phi, x) := \bar{s}_1(x) \frac{\phi - \min(\phi, gB(x))}{g} + \bar{s}_2(x) \left(\frac{\min(\phi, gB(x))}{g} - S(x) \right) \quad \text{and} \quad V_k^n = V(\phi_k^n, x_k).$$

Here we describe the function used to compute our solution, the explanation of how we use them are in the section §2.3.3. The discrete potential is computed based on the two mass equations of (*SW/DF*), i.e.

$$\mathcal{S}_k^n(\phi) := V(\phi, x) - V_k^n + \frac{\delta_t}{\delta_k} \sum_{f \in \mathbb{F}_k} (F_{1,f,k}(\phi, x) + F_{2,f,k}(\phi, x)) \cdot N_k^{k_f} \mu_f^k, \quad (19)$$

with the function $F_{1,f,k}$ an approximation of the surface flux $V_1 \bar{\mathbf{u}}_1$ and the function $F_{2,f,k}$ an approximation of the underground flux $-\kappa V_2 \nabla \phi$ at the face f along the time step. The flux for the free surface layer is defined by,

$$F_{1,f,k}(\phi, x) = (V_1^{n*}(\phi, x) \bar{\mathbf{u}}^n)_f - \gamma \delta_t \left(\frac{V_1^{n*}(\phi, x)}{\delta} \right)_f [\phi]_k^{k_f} N_k^{k_f}, \quad (20)$$

with $\gamma \in \mathbb{R}$ a given regularization parameter more thoroughly describe in [7] and $\bar{\mathbf{u}}$ the given velocity of free surface. The function $V_{1,k}^{n*}$ is an approximation of the volume of water in one layer, without the mass exchange term,

$$V_{1,k}^{n*}(\phi, x) := V_{1,k}^n - \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} F_{1,f,k}(\phi, x) \cdot N_k^{k_f} \mu_f^k. \quad (21)$$

The flux for the underground layer is defined by

$$F_{2,f,k}(\phi, x) = -\kappa_f \left(\frac{V_2(V(\phi, x), x)}{\delta} \right)_f [\phi]_k^{k_f} N_k^{k_f}, \quad (22)$$

with κ_f an approximation of κ at the face f .

For readability, we note $F_{1,f,k}^{n+1} = F_{1,f,k}(\phi_k^{n+1}, x_k)$, $F_{2,f,k}^{n+1} = F_{2,f,k}(\phi_k^{n+1}, x_k)$. The volume for the free surface and the underground layer are calculated with (7) and note $V_{1,k}^{n+1} = V_1(V(\phi_k^{n+1}, x_k), x_k)$ and $V_{2,k}^{n+1} = V_2(V(\phi_k^{n+1}, x_k), x_k)$.

To compute the velocity in the surface layer at the new time step $\bar{\mathbf{u}}_k^{n+1}$. We set

$$\begin{aligned} \alpha_k^{n+1} \bar{\mathbf{u}}_k^{n+1} &= V_{1,k}^n \bar{\mathbf{u}}_k^n - \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(\bar{\mathbf{u}}_k^n \left(F_{1,f,k}^{n+1} \cdot N_k^{k_f} \right)^+ - \bar{\mathbf{u}}_{k_f}^n \left(F_{1,f,k}^{n+1} \cdot N_k^{k_f} \right)^- \right) \mu_f^k \\ &\quad - V_{1,k}^{n*} \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} (\phi^{n+1})_f N_k^{k_f} \mu_f^k \end{aligned} \quad (23)$$

with the mass coefficient $\alpha_k^{n+1} = V_{1,k}^{n+1} + \delta_t^{n+1} \frac{G_k^{n+1}}{2}$ and the positive and negative part functions define by $(a)^\pm = \frac{|a| \pm a}{2}$.

Note that at this stage, the mass exchanged G_k^{n+1} is still unknown.

Thus, the mass exchange is obtained by considering the difference between the volume of water without the mass exchange $V_{1,k}^{n*}$, and with the mass exchange $V_{1,k}^{n+1}$. It reads

$$\delta_t^{n+1} G_k^{n+1} := V_{1,k}^{n*} - V_{1,k}^{n+1} \quad \text{hence} \quad \alpha_k^{n+1} = \frac{V_{1,k}^{n+1} + V_{1,k}^{n*}}{2}. \quad (24)$$

For the remainder of this article, $V_{1,k}^{n*}$ would represent the function (21) evaluated at the point $(\phi_k^{n+1}, x_k) \in \mathbb{R}^2$. To ensure the scheme to be entropy satisfying, the time step must satisfy the following implicit CFL condition, see §2.3.2

$$\frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(F_{1,f,k}^{n+1} \cdot N_k^{k_f} \right)^- \mu_f^k < \frac{V_{1,k}^{n*}}{2}. \quad (25)$$

2.3.2 Numerical analysis

The current section is devoted to the analysis of the numerical scheme presented in §2.3.1. More precisely, we proof that the scheme is entropy satisfying, considering the mechanic energy as the mathematical entropy.

Theorem 4. *Assuming the CFL condition (25) is satisfied and $\gamma \geq 1$, the discrete mechanical energy satisfies the following dissipation law*

$$E_k^{n+1} + \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(\mathcal{G}_{\mathcal{E},f}^{n+1} + \mathcal{G}_{\mathcal{K},f}^{n+1} \right) \cdot N_k^{k_f} \mu_f^k \leq E_k^n - \delta_t \mathcal{P}_k^{n+1}$$

with the discrete mechanical energy $E_k^n = \mathcal{E}_k^n + \mathcal{K}_k^n$, the discrete potential energy $\mathcal{E}_k^n = \mathcal{E}(V_k^n, x_k)$ defined in (14) and the discrete kinetic energy $\mathcal{K}_k^n = V_{1,k}^n \frac{|\bar{\mathbf{u}}_k^n|^2}{2}$. The discrete flux of potential energy reads

$$\mathcal{G}_{\mathcal{E},f}^{n+1} := (\phi^{n+1})_f \left(F_{1,f}^{n+1} + F_{2,f}^{n+1} \right) - [\phi^{n+1}]_k^{k_f} [V^{n+1} \bar{\mathbf{u}}^n]_k^{k_f},$$

the discrete flux of kinetic energy reads

$$\mathcal{G}_{\mathcal{K},f}^{n+1} \cdot N_k^{k_f} := \frac{|\bar{\mathbf{u}}_k^n|^2}{2} \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^+ - \frac{|\bar{\mathbf{u}}_{k_f}^n|^2}{2} \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^- + \delta_t^{n+1} \left[\frac{V_1^{n*}}{\delta} \right]_k^{k_f} \left| [\phi^{n+1}]_k^{k_f} \right|^2.$$

and the discrete dissipation due to the drag force reads

$$\mathcal{P}_k^{n+1} := \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(\kappa_f \left(\frac{V_2^{n+1}}{\delta} \right)_f + \gamma \delta_t^{n+1} \left(\frac{V^{n*}}{\delta} \right)_f \right) \left| [\phi^{n+1}]_k^{k_f} \right|^2$$

Proof. The estimate of Theorem 4 is a corollary of the estimates of Lemma 1 and Lemma 2. \square

Corollary 1. *The inequality of Theorem 4 is a dissipation law in the sense that, by summing over the whole mesh, the global mechanical energy is decreasing, i.e.*

$$\sum_{k \in \mathbb{T}} E_k^{n+1} \leq \sum_{k \in \mathbb{T}} E_k^n.$$

Proof. The results come from the fact the discrete fluxes are anti-symmetric and the dissipation term due to the drag force is non-negative. \square

Lemma 1 (Potential energy estimate). *The scheme (19) with the fluxes (20) and (22) satisfy the following potential energy inequality*

$$\mathcal{E}_k^{n+1} + \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \mathcal{G}_{\mathcal{E},f}^{n+1} \cdot N_f^k \mu_f^k \leq \mathcal{E}_k^n + \delta_t^{n+1} \mathcal{Q}_k^{n+1} + \mathcal{P}_k^{n+1},$$

with the discrete work of the conservative forces and the discretization error

$$\mathcal{Q}_k^{n+1} := \frac{V_{1,k}^{n*} \bar{\mathbf{u}}_k^n}{\delta_k} \cdot \sum_{f \in \mathbb{F}_k} [\phi^{n+1}]_k^{k_f} N_k^{k_f} \mu_f^k$$

Proof. The function \mathcal{E} is convex in V by parts on the intervals $I_2(x) = [0, \bar{V}_2(x)]$ and $I_1(x) = [\bar{V}_2(x), +\infty[$. We remind that $\phi = g\left(\frac{V_1}{s_1} + \frac{V_2}{s_2} + S\right)$. In the case where V_k^{n+1} and V_k^n are both contain in $I_1^k = I_1(x_k)$ (resp. $I_2^k = I_2(x_k)$), the energy potential function is convex and we can use the same proof find in [7]. In the case where $V_k^{n+1} \in I_1^k$ and $V_k^n \in I_2^k$ (resp. $V_k^{n+1} \in I_2^k$ and $V_k^n \in I_1^k$), the energy potential function \mathcal{E} isn't convex on the interval $[V_k^n, V_k^{n+1}]$.

The function \mathcal{E} is continuous by part on $[0, \bar{V}_2[$ and $]\bar{V}_2, +\infty[$; the limit on the left hand side and on the right hand side of \bar{V}_2 of $\mathcal{E}(V, x)$ is equal to $g\bar{V}_2\left(\frac{B+S}{2}\right)$ for both of them so the function \mathcal{E} is continuous on $[0, +\infty[$.

We can now use the Taylor's theorem on two intervals $[V_k^n, \bar{V}_2]$ and $[\bar{V}_2, V_k^{n+1}]$ with the mean-value forms of the remainder

$$\mathcal{E}_k^n = \mathcal{E}(\bar{V}_{2,k}, x_k) + (V_k^n - \bar{V}_{2,k}) \bar{\phi}_k + \frac{|V_k^n - \bar{V}_{2,k}|^2}{2} \partial_{V^2} \mathcal{E}(\tilde{V}_k, x_k) \quad (26)$$

$$\mathcal{E}(\bar{V}_{2,k}, x_k) = \mathcal{E}_k^{n+1} + (\bar{V}_{2,k} - V_k^{n+1}) \phi_k^{n+1} + \frac{|\bar{V}_{2,k} - V_k^{n+1}|^2}{2} \partial_{V^2} \mathcal{E}(\hat{V}_k, x_k) \quad (27)$$

with $\phi_k^{n+1} = \partial_V \mathcal{E}(V_k^{n+1}, x_k)$ by construction of the function \mathcal{E} , $\tilde{V}_k \in]V_k^n, \bar{V}_{2,k}[$ and $\hat{V}_k \in]\bar{V}_{2,k}, V_k^{n+1}[$. We can also note that

$$(V_k^{n+1} - V_k^n) \phi_k^{n+1} = -\phi_k^{n+1} (\bar{V}_{2,k} - V_k^{n+1}) - \bar{\phi}_k (V_k^n - \bar{V}_{2,k}) + (\phi_k^{n+1} - \bar{\phi}_k) (\bar{V}_{2,k} - V_k^n) \quad (28)$$

Replacing $\mathcal{E}(\bar{V}_{2,k}, x_k)$ in (26) by (27) and then using the expression (28), we got

$$\mathcal{E}_k^{n+1} - \mathcal{E}_k^n = (V_k^{n+1} - V_k^n) \phi_k^{n+1} - (\phi_k^{n+1} - \bar{\phi}_k) (\bar{V}_{2,k} - V_k^n) - \frac{|V_k^n - \bar{V}_{2,k}|^2}{2} \partial_{V^2} \mathcal{E}(\tilde{V}_k, x_k) - \frac{|\bar{V}_{2,k} - V_k^{n+1}|^2}{2} \partial_{V^2} \mathcal{E}(\hat{V}_k, x_k)$$

The function \mathcal{E} is convex by part on the intervals I_1^k and I_2^k so the second derivative term of the equations are positive. The term $(\phi_k^{n+1} - \bar{\phi}_k) (\bar{V}_{2,k} - V_k^n)$ is positive by monotonic increasing of the function ϕ and $V_k^n \leq \bar{V}_{2,k}$ (resp. $V_k^n \geq \bar{V}_{2,k}$).

Recall that $a_k = (a)_f - [a]_k^{k_f}$, using the discrete potential equation (19) and dividing the potential between the symmetric and anti-symmetric part, it implies the following inequality

$$\begin{aligned} \mathcal{E}_k^{n+1} - \mathcal{E}_k^n &\leq -\frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(F_{1,f,k}^{n+1} + F_{2,f,k}^{n+1} \right) (\phi^{n+1})_f \cdot N_k^{k_f} \mu_f^k \\ &\quad + \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(V_{1,k}^{n*} \bar{\mathbf{u}}_k^n + [V_1^{n*} \bar{\mathbf{u}}^n]_k^{k_f} - \gamma \delta_t^{n+1} \left(\frac{V_1^{n*}}{\delta} \right)_f [\phi^{n+1}]_k^{k_f} N_k^{k_f} + F_{2,f,k}^{n+1} \right) \cdot N_k^{k_f} [\phi^{n+1}] \mu_f^k \end{aligned}$$

From this inequality we recognize the discrete work of the force \mathcal{Q}_k^{n+1} and the source term \mathcal{P}_k^{n+1} in the second term. To find the anti-symmetrical flux $\mathcal{G}_{\mathcal{E},f,k}^{n+1}$, you simply add the sum term $[V_1^{n*} \bar{\mathbf{u}}^n]_k^{k_f} \cdot N_k^{k_f} [\phi^{n+1}]$ to the first term of the inequality.

$$\mathcal{E}_k^{n+1} - \mathcal{E}_k^n \leq -\frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \mathcal{G}_{\mathcal{E},f,k}^{n+1} \cdot N_k^{k_f} \mu_f^k + \delta_t^{n+1} \mathcal{Q}_k^{n+1} + \mathcal{P}_k^{n+1}$$

With this inequality, we can conclude the result of the lemma. \square

Lemma 2 (Kinetic energy estimate). *Assuming the CFL condition (25) is satisfied, then the numerical scheme (19), (23) and (21) satisfies a discrete kinetic energy inequality such as*

$$\mathcal{K}_k^{n+1} + \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \mathcal{G}_{\mathcal{K},f}^{n+1} \cdot N_k^{k_f} \mu_f^k \leq \mathcal{K}_k^n - \delta_t^{n+1} (\mathcal{Q}_k^{n+1} - \mathcal{R}_k^{n+1}).$$

with \mathcal{R}_k^{n+1} the discretization error

$$\mathcal{R}_k^{n+1} := \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(\frac{V_1^{n*}}{\delta} \right)_f \left| [\phi^{n+1}]_k^{k_f} \right|^2 \mu_f^k.$$

Proof. We write the equation (31) under the non-conservative form, using (21) to replace $V_{1,k}^n$ and the fact that $a = (a)^+ - (a)^-$ for all function a , we write

$$\bar{\mathbf{u}}_k^{n*} = \bar{\mathbf{u}}_k^n + \frac{\delta_t}{\delta_k} \sum_{f \in \mathbb{F}_k} \frac{\bar{\mathbf{u}}_{k_f}^n - \bar{\mathbf{u}}_k^n}{V_{1,k}^{n*}} \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^- \mu_f^k - \frac{\delta_t}{\delta_k} \sum_{f \in \mathbb{F}_k} (\phi^{n+1})_f N_k^{k_f} \mu_f^k \quad (29)$$

The last term can be rewrite as

$$\sum_{f \in F_k} (\phi^{n+1})_f N_k^{k_f} \mu_f^k = \sum_{f \in F_k} [\phi^{n+1}]_k^{k_f} N_k^{k_f} \mu_f^k.$$

We recall the equality $a(b-a) = \frac{|b|^2}{2} - \frac{|a|^2}{2} - \frac{|b-a|^2}{2}$, with $a, b \in \mathbb{R}$. To obtain the discrete kinetic energy \mathcal{K}_k^{n*} , we multiply (23) by $V_{1,k}^{n*} \bar{\mathbf{u}}_k^n$

$$\mathcal{K}_k^{n*} - V_{1,k}^{n*} \frac{|\bar{\mathbf{u}}_k^n|^2}{2} = \frac{\delta_t}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(\frac{|\bar{\mathbf{u}}_{k_f}^n|^2}{2} - \frac{|\bar{\mathbf{u}}_k^n|^2}{2} \right) \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^- \mu_f^k - \delta_t \mathcal{Q}_k^{n+1} + S_k^{n+1}$$

with the source term

$$S_k^{n+1} = \frac{|\bar{\mathbf{u}}_k^{n*} - \bar{\mathbf{u}}_k^n|^2}{2} V_{1,k}^{n*} - \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \frac{|\bar{\mathbf{u}}_{k_f}^n - \bar{\mathbf{u}}_k^n|^2}{2} \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^- \mu_f^k.$$

We use (21) again and replace $V_{1,k}^{n*}$

$$\mathcal{K}_k^{n*} - \mathcal{K}_k^n = -\frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(\frac{|\bar{\mathbf{u}}_k^n|^2}{2} \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^+ - \frac{|\bar{\mathbf{u}}_{k_f}^n|^2}{2} \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^- \right) \mu_f^k - \delta_t \mathcal{Q}_k^{n+1} + S_k^{n+1}.$$

It remains to estimate an upper bound of the last term using (29) and the Jensen's inequality, we get

$$\begin{aligned} S_k^{n+1} &\leq \left| \frac{\delta_t^{n+1}}{\delta_k} \right|^2 \sum_{f \in \mathbb{F}_k} \left| [\phi^{n+1}]_k^{k_f} \right|^2 \mu_f^k \\ &\quad + \frac{4}{V_{1,k}^{n*}} \left| \frac{\delta_t^{n+1}}{\delta_k} \right|^2 \sum_{f \in \mathbb{F}_k} \left| \left([\bar{\mathbf{u}}^n]_k^{k_f} \sqrt{\left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^-} \right) \left(\sqrt{\left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^-} \right) \right|^2 \mu_f^k \\ &\quad - 2 \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left| [\bar{\mathbf{u}}^n]_k^{k_f} \right|^2 \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^- \mu_f^k. \end{aligned}$$

Then the first term of the inequality is separated into the symmetric and antisymmetric part around each face and the second term is bounded under a form of the third term using the Cauchy-Schwarz inequality. We get

$$\begin{aligned} S_k^{n+1} &\leq \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \delta_t^{n+1} \left(\left(\frac{V_{1,k}^{n*}}{\delta} \right)_f - \left[\frac{V_{1,k}^{n*}}{\delta} \right]_k^{k_f} \right) \left| [\phi^{n+1}]_k^{k_f} \right|^2 \mu_f^k \\ &\quad - \frac{4\delta_t^{n+1}}{V_{1,k}^{n*} \delta_k} \left(\sum_{f \in \mathbb{F}_k} \left| [\bar{\mathbf{u}}^n]_k^{k_f} \right|^2 \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^- \mu_f^k \right) \left(\frac{V_{1,k}^{n*}}{2} - \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^- \mu_f^k \right) \end{aligned}$$

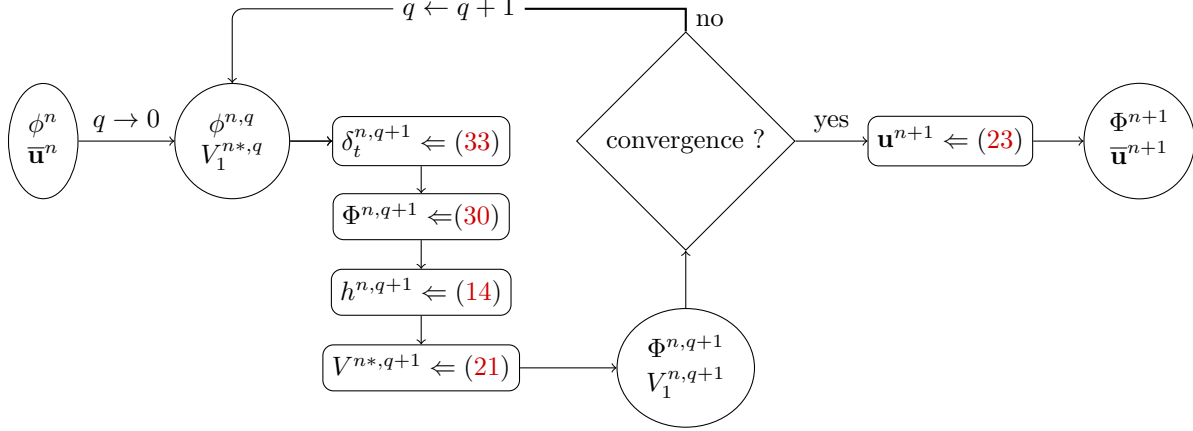
Under the CFL condition (25), the last term is non-negative. Each parts of the first term of the inequality are respectively place in the discretization error $\mathcal{R}_{\mathcal{K},f}^{n+1}$ and in the flux $\mathcal{G}_{\mathcal{K},f}^{n+1}$, hence we get the estimate

$$\mathcal{K}_k^{n*} + \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \mathcal{G}_{\mathcal{K},f}^{n+1} \cdot N_k^{k_f} \mu_f^k \leq \mathcal{K}_k^n + \delta_t^{n+1} \left(\mathcal{R}_f^{n+1} - \mathcal{Q}_k^{n+1} \right).$$

Eventually, it remains to estimate the energy balance of the second step of the scheme (32). Using the (24), we get

$$\bar{\mathbf{u}}_k^{n+1} - \bar{\mathbf{u}}_k^{n*} = \delta_t^{n+1} \frac{G_k^{n+1}}{2V_{1,k}^{n*}} \bar{\mathbf{u}}_k^{n+1}.$$

Figure 2: Representation of the algorithm



Then multiplying by $\bar{\mathbf{u}}_k^{n+1}$ and using the equality $a(a-b) = \frac{a^2}{2} - \frac{b^2}{2} + \frac{|b-a|^2}{2}$ we get

$$V_{1,k}^{n*} \frac{|\bar{\mathbf{u}}_k^{n+1}|^2}{2} - V_{1,k}^{n*} \frac{|\bar{\mathbf{u}}_k^{n*}|^2}{2} = \delta_t^{n+1} \frac{G_k^{n+1}}{2} |\bar{\mathbf{u}}_k^{n+1}|^2 - V_{1,k}^{n*} \frac{|\bar{\mathbf{u}}_k^{n+1} - \bar{\mathbf{u}}_k^{n*}|^2}{2}.$$

Using (24) one more time we get the estimate

$$\mathcal{K}_k^{n+1} - \mathcal{K}_k^{n*} = -V_{1,k}^{n*} \frac{|\bar{\mathbf{u}}_k^{n+1} - \bar{\mathbf{u}}_k^{n*}|^2}{2} \leq 0.$$

We conclude by composing the two estimates. \square

2.3.3 Practical computation

Here we explain how to use the function describe in §2.3.1 to compute the unknowns ϕ_k^{n+1} and $\bar{\mathbf{u}}_k^{n+1}$. The Figure 2 show the algorithm used. This scheme is an implicit-explicit non-linear method using finite volume. First we use an Newton method on (19) and then an explicit scheme (23) to calculate the velocity. At each iteration of the Newton method we also calculate the time step δ_t^{n+1} control by a CFL condition to keep the energy conservation.

Newton method for the implicit part. To calculate our discrete total mass equation we want to find ϕ_k^{n+1} , for each volume control k , such that $\mathcal{S}_k^n(\phi_k^{n+1}) = 0$. This equation is non-linear, thus we need to use a iterative process called the Newton method, to calculate the solution. Each step of the Newton is denote q and for visibility, we define the vector of all $\phi_k^{n,q}$ as $\Phi^{n,q} = \{\phi_k^{n,q}\}_{k \in \mathbb{T}}$.

$$\begin{cases} J_{\mathcal{S}}(\Phi^{n,q}) \delta_{\Phi}^{n,q+1} = -\mathcal{S}_k^n(\Phi^{q,n}) \\ \Phi^{n,q+1} = \Phi^{n,q} + \delta_{\Phi}^{n,q+1} \end{cases} \quad (30)$$

with $J_{\mathcal{S}}$ an approximation of the Jacobian of \mathcal{S} .

As show in the left part of the Figure 2, we consider that we have all the values we need at the time n, q to compute the time $n, q+1$. We start by solving the system (30) to compute the solution $\Phi^{n,q+1}$. Then we calculate the height of the water with (14) to obtain the boundary condition of $\phi_k^{n,q+1}$. We denote $h_k^{n,q+1} = h(V(\phi_k^{n,q+1}, x_k), x_k)$. With the boundary set for $\phi_k^{n,q+1}$, we used (21) and define $V_{1,k}^{n*,q+1} = V_{1,k}^{n*}(\phi_k^{n*,q}, x_k)$. We then compute $\delta_t^{n,q+1}$ with the equation (33) explained in the paragraph about CFL in this section.

When $|\mathcal{S}(\phi_k^{q,n+1})| < \epsilon$, with $\epsilon \ll 1$ then $\phi_k^{n+1} = \phi_k^{n,q+1}$ and we find the solution to our problem. If not we iterate another step of the Newton method.

If the underground volume of water is zero then there is no system (*SW/DF*), because of the hypothesis on the saturation of the medium, if $V_{2,k}^n = 0$ then $V_{1,k}^n = 0$. In the case of $V_{1,k}^n = 0$, then the momentum equation vanish and it remain only the non-linear system (30) to solve implicitly. Then we fix a δ_t^n maximum so that we don't end the computation with only one time step, because the CFL isn't necessary anymore.

Upwind finite volume for the explicit part. The scheme for the velocity computation is obtained using an upwind scheme for the horizontal advection term, a centered estimation of the potential source term and an implicit centered scheme for the vertical advection term, corresponding to $\lambda = 0$ in the velocity at the surface (16). The motivation of the last choice is to minimize the dissipation of energy see Theorem 6.

Practically, the velocity is calculate in two steps, one at the intermediary time without the mass exchange term and then with it. The following equations is an approximation of the momentum equation (*SW/DF*) without the mass exchange term $\mathbf{u}_B G$.

$$V_{1,k}^{n*} \bar{\mathbf{u}}_k^{n*} = V_{1,k}^n \bar{\mathbf{u}}_k^n - \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} \left(\bar{\mathbf{u}}_k^n \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^+ - \bar{\mathbf{u}}_{k_f}^n \left(F_{1,f}^{n+1} \cdot N_k^{k_f} \right)^- \right) \mu_f^k - V_{1,k}^{n*} \frac{\delta_t^{n+1}}{\delta_k} \sum_{f \in \mathbb{F}_k} (\phi^{n+1})_f N_k^{k_f} \mu_f^k \quad (31)$$

The we add the mass exchange term and find the unknown $\bar{\mathbf{u}}_k^{n+1}$ with

$$\left(V_{1,k}^{n+1} + \delta_t^{n+1} G_k^{n+1} \right) \bar{\mathbf{u}}_k^{n+1} = V_{1,k}^{n*} \bar{\mathbf{u}}_k^{n*} \quad (32)$$

with the expression for G_k^{n+1} known (24). The discrete velocity $\bar{\mathbf{u}}_k^{n+1}$ is not defined if both $V_{1,k}^{n+1}$ and $V_{1,k}^{n*}$ are zero. In order to avoid this problem, as it is classically done in shallow water flows, we set $\bar{\mathbf{u}}_k^{n+1} = 0$ when $V_{1,k}^{n+1} = 0$.

CFL used in practice. The CFL condition (25) is implicit and we would like to compare it with a more classical CFL condition ; thus this CFL condition would be more restrictive. The numerical flux $F_{1,f}^{n+1}$ can be bound as follow

$$\left| F_{1,f}^{n+1} \cdot N_k^{k_f} \right| \leq \left| (V_1^{n*} \bar{\mathbf{u}}^n)_f \cdot N_k^{k_f} \right| + \gamma \delta_t^{n+1} \left(\frac{V_1^{n*}}{\delta} \right)_f \left| [\phi^{n+1}]_k^{k_f} \right|.$$

The sum over the face are normalized, so for each face f can be lower bound by the right hand side of (25),

$$\frac{1}{2} > \frac{\delta_t^{n+1}}{\min(\delta_k, \delta_{k_f})} \frac{\left| (V_1^{n*} \bar{\mathbf{u}}^n)_f \cdot N_k^{k_f} \right|}{\min(V_{1,k}^{n*}, V_{1,k_f}^{n*})} + \gamma \left| \frac{\delta_t^{n+1}}{\min(\delta_k, \delta_{k_f})} \right|^2 \frac{(V_1^{n*})_f \left| [\phi^{n+1}]_k^{k_f} \right|}{\min(V_{1,k}^{n*}, V_{1,k_f}^{n*})}.$$

Thanks to this inequality, we have the polynomial over $\frac{\delta_t}{\min(\delta_k, \delta_{k_f})}$

$$\left(x \frac{\delta_t}{\min(\delta_k, \delta_{k_f})} \right)^2 + y \frac{\delta_t}{\min(\delta_k, \delta_{k_f})} - 1,$$

with $x = \frac{\sqrt{2\gamma(V_1^{n*})_f \left| [\phi^{n+1}]_k^{k_f} \right|}}{\min(V_{1,k}^{n*}, V_{1,k_f}^{n*})}$ and $y = 2 \frac{\left| (V_1^{n*} \bar{\mathbf{u}}^n)_f \cdot N_k^{k_f} \right|}{\min(V_{1,k}^{n*}, V_{1,k_f}^{n*})}$, the two quantities x and y are non-negative so the polynomial is negative when $(x + y) \frac{\delta_t}{\min(\delta_k, \delta_{k_f})} < 1$ and we can now compute our new CFL condition

$$\frac{\left| (V_1^{n*} \bar{\mathbf{u}}^n)_f \cdot N_k^{k_f} \right| + \sqrt{\frac{\gamma}{2} (V_1^{n*})_f \left| [\phi^{n+1}]_k^{k_f} \right|}}{\min(V_{1,k}^{n*}, V_{1,k_f}^{n*})} \frac{\delta_t}{\min(\delta_k, \delta_{k_f})} \leq \frac{1}{2}. \quad (33)$$

2.4 Numerical experiments

The purpose of this section is to numerically illustrate the result of our model and numerical method in 1D on a homogeneous grid, i.e $\delta_k = \delta_x \in \mathbb{R}$. The gravity acceleration is going to be $g = 9.81$. The length of the channel is $L = 1$. The other parameters are bound to change between different cases, the regularization parameter γ need to be greater than 1 but depends on whether there are shocks or not in the simulations. The porosity \bar{s}_i and the hydraulic conductivity $\bar{\kappa}$ depends of the type of porous media we are studying.

In §2.4.1 and §2.4.2 we show the asymptotic behavior of our scheme: a case with an analytical stationary Dupuit-Forchheimer solution and a analytical stationary solution of a transcritical steady flow. It is also worth testing the behavior of our model based on the flow distribution between the porous layer and the free surface.

2.4.1 Steady flow between two reservoirs

For any smooth solutions such as the water table is always under the bathymetry, we can find a steady states solutions such as $\partial_t h = 0$ in (DF) in the case where the substratum S is constant.

$$\bar{\kappa} \nabla \cdot (h \nabla (h + S)) = 0. \quad (34)$$

The porous layer height is $B(x) = 5.0$, the water height at the entrance is fixed at $h(t, 0) = 0.5$ and the one at the exit at $h(t, L) = 4.5$. The analytical solution for (34) is

$$h(t, x) = \sqrt{(h(t, L)^2 - h(t, 0)^2)x + h(t, 0)^2}.$$

The numerical solution with $\bar{\kappa} = 1.d - 3$ and porosity at $\bar{s}_2 = 0.6$ are plotted together with the analytical solution for a regular space step $\delta_x = 10^{-3}$ and an initial condition $h(0, x) = -4x + 4.5$ in Figure ?.

The goal here is to verify the behavior of our scheme when the shallow water part of the model isn't interfering. The simulation is performed on multiple space step $\delta_x = \{1/200, 1/500, 1/1000, 1/5000\}$ then we plot the L^2 -errors of the water height between the analytical solution and the numerical one.

2.4.2 Transcritical steady flow with constrained hydraulic jump

For any smooth substratum and bathymetry choose such as $B(x) = S(x)$

$$K = h(x) + S(x) + \frac{H_c^3}{2h^2(x)}$$

with $H_c = \left(\frac{Q}{\sqrt{g}}\right)^{\frac{2}{3}}$ the critical water depth.

The water depth is solution of this third order polynomial

$$h^3 + (B - S)h^2 + \frac{H_c^3}{2} = 0$$

2.4.3 Groundwater overflowing

This simulation is an analysis of the scheme behavior when the Dupuit-Forchheimer is prevailing, more precisely how it impacts the free surface flow. For this test case and the following one, the substratum is fixed at zero in the whole space, and the bathymetry is $2x$. The water height at the beginning is constant $h(0, x) = 0.5$, the left boundary condition is a wall and the height is fixed for the right boundary condition with

$$h(t, 1) = 1.2 + 0.3 \frac{\cos(\pi t)}{T/2}.$$

In this situation, all the added water of this simulation come from the porous media. The porosity is chosen such as the underground layer hold as much water than it release water, $\bar{s}_2 = 0.5$ and the hydraulic conductivity is equivalent to a very fine sand layer, $\bar{\kappa} = 10^{-5}$.

The time of the simulation is long $T > 1000$, the space interval is $\delta_x = 5.10^{-3}$ and the CPR-parameter is $\gamma = 1$; there is no shock in this simulation so the parameter doesn't need to be increased.

2.4.4 Waves behavior on a sand beach

The previous simulation was showing the behavior of the Dupuit-Forchheimer part of the scheme, this one is for the shallow water part. The case is the common waves behavior on a sand beach case but this time, instead of considering a non-permeable bathymetry, the beach is a porous media. The settings are the same than before for the substratum and the bathymetry. The left and right boundary conditions are a wall, the initial condition is a Riemann problem such as

$$h(0, x) = \begin{cases} 1.8 & \text{if } x < 0.5 \\ 1.1 & \text{if not} \end{cases}$$

and $\bar{\mathbf{u}}(0, x) = 0$.

This simulation is going to generate shock so the CPR-parameter is increased $\gamma = 10$, it last for a long time again $T > 1000$ and the space interval is of $\delta_x = 10^{-3}$ The hydraulic conductivity is the same than before but this time, the porosity is lesser $\bar{s}_2 = 0.3$; it means the pores are smaller than before.

2.5 Congested formulation for confined aquifer

Our approach until now didn't take into account underground river or confined aquifer situations. To add a congested part to the (*SW/DF*) model, we are using the work done in [5] with a congested shallow water model. We chose this approach because it used the same numerical scheme and because our model is derived from a shallow water model.

If the case is congested, it means we have a non permeable roof R dependent on the space x above the interface between the free surface and the porous media such that $R(x) \geq B(x) \geq S(x)$. The potential is rewrite

$$\phi(V, x) = g(h(V, x) + S(x)) + \frac{g}{\beta^2} \frac{(V - \bar{V}(x))}{\bar{s}_1(x)}$$

with \bar{V} is the maximum value of the water height between the substratum and the roof and β the coefficient the relaxation parameter.

The formulation of (*SW/DF*) doesn't change however we need to update the potential energy equation (14).

$$\mathcal{E}(V, x) = g \left(V_1(V, x) \left(\frac{V_1(V, x)}{2\bar{s}_1(x)} + B(x) \right) + V_2(V, x) \left(\frac{V_2(V, x)}{2\bar{s}_2(x)} + S(x) \right) \right) + g \frac{(V - \bar{V}(x))_+^2}{2\beta^2}$$

This new ϕ implies that the computation of volume of the water in the numerical scheme §2.3.1 also change. It now reads

$$V(\phi, x) = \bar{s}_2(x) \left(\frac{\min(\phi, gB(x))}{g} - S(x) \right) + \bar{s}_1(x) \left(\frac{\min(\phi, gR(x)) - \min(\phi, gB(x))}{g} \right) + \frac{\beta^2}{1 + \beta^2} \bar{s}_1(x) \left(\frac{\max(\phi, gR(x))}{g} - R(x) \right)$$

3 Conclusion

References

- [1] AUDUSSE, EMMANUEL, BRISTEAU, MARIE-ODILE, PERTHAME, BENOÎT, AND SAINTE-MARIE, JACQUES, A multilayer saint-venant system with mass exchanges for shallow water flows. derivation and numerical validation, ESAIM: M2AN, 45 (2011), pp. 169–200.
- [2] A. J.-C. DE SAINT-VENANT ET AL., Théorie du mouvement non-permanent des eaux, avec application aux crues des rivières et à l'introduction des marées dans leur lit, CR Acad. Sci. Paris, 73 (1871), p. 5.
- [3] J. DUPUIT, Études théoriques et pratiques sur le mouvement des eaux dans les canaux découverts et à travers les terrains perméables: avec des considérations relatives au régime des grandes eaux, au débouché à leur donner, et à la marche des alluvions dans les rivières à fond mobile, 1863.
- [4] J.-F. GERBEAU AND B. PERTHAME, Derivation of viscous Saint-Venant system for laminar shallow water; numerical validation, PhD thesis, INRIA, 2000.
- [5] E. GODLEWSKI, M. PARISOT, J. SAINTE-MARIE, AND F. WAHL, Congested shallow water model: roof modelling in free surface flow, ESAIM: Mathematical Modelling and Numerical Analysis, 52 (2018), pp. 1679 – 1707.
- [6] ———, Congested shallow water model: on floating body, SMAI Journal of Computational Mathematics, (2021).
- [7] M. PARISOT AND J.-P. VILA, Centered-potential regularization for the advection upstream splitting method, SIAM Journal on Numerical Analysis, (2016).