*Monografías de la Real Academia de Ciencias de Zaragoza* **31**, 89–119, (2008).

# The shallow water equations: An example of hyperbolic system

P. García-Navarro, P. Brufau, J. Burguete and J. Murillo,

Fluid Mechanics. CPS. University of Zaragoza

50018 Zaragoza, Spain

**Abstract**

Many problems of river management and civil protection consist of the evaluation of the maximum water levels and discharges that may be attained at particular locations during the development of an exceptional meteorological event. There is also the prevision of the scenario subsequent to the almost instantaneous release of a great volume of liquid. The situation is that of the breaking of a man made dam. There is therefore a necessity to develop adequate numerical models able to reproduce situations originated by the irregularities of a non-prismatic bed. It is also necessary to trace their applicability considering the difficulty of developing a model capable of producing solutions of the complete equations despite the irregular character of the river bed. When trying to use mathematical models as a predictive tool in the simulation of free surface flows, the hypothesis of one-dimensional models are not always valid. Such is the case when dealing with compound, or highly irregular, cross-section configurations, abrupt contractions and expansions, or rivers of high curvature. When trying to reproduce these hydraulic situations, it becomes necessary to use a two-dimensional formalism which takes into consideration the influence of transverse components of the flow. Many efforts have been recently devoted to the development of multidimensional techniques for free surface flows.

## 1 Introduction

Free surface flows common in Hydraulics are usually described by means of the shallow water equations, provided that the representative vertical dimensions are small with respect to the horizontal dimensions. Despite their simplicity, this description is valid in many practical applications, rendering worthwhile the efforts in developing good numerical methods to solve the corresponding system of differential equations.

As in any other case of integration of a system of non linear partial differential equations, the first difficulty is the choice of the numerical scheme. It must be pointed out that numerical results will be a better/worse approach of the exact solution of the equations depending on the technique adopted for their discretization and on the computational tools used. Moreover, even using the exact solution of the governing equations, the model predictions can differ from the observation if the mathematical model does not contain all the relevant physics, that is, when the model hypothesis are not valid.

Numerical modelling is an option versus experimental techniques. It is sometimes preferred mainly due to economical reasons but also because of the amount of accesible information it provides. They are actually two different and complementary tools since, in particular, experimental data are necessary in order to validate numerical results. Numerical methods are nowadays a common tool to predict flow properties both for steady and unsteady situations of practical interest in Hydraulics. The application of finite differences and finite volumes has been widely reported in particular. We will be concerned with numerical methods from the finite volume point of view for the resolution of shallow water equations in one-dimensional (1D) and two-dimensional (2D) approaches. The 1D approach is valid whenever a more detailed description is not necessary. In other cases of very irregular geometry, sudden expansions or contractions, high curvature, etc, this approach can be insufficient and a 2D or 3D description involving other velocity components becomes necesary.

Traditionally, 1D schemes have been based on central differences and they still dominate the commercial software in the field of computational hydraulics [2, 1] These schemes were not developed taking into account the mathematical structure of the system of equations. Hence the quality of their solution depends on the flow conditions. leading in some cases to totally incorrect results. Second order schemes are known for their oscillatory behaviour in presence of strong gradients [3] and, therefore, to require some kind of artificial viscosity used as the easiest way to get rid of undesired numerical effects. The disadvantage of this option is the global character, that can affect regions of gradually varied flow, and the necessary tunning.

More recent schemes, known as high resolution schemes, pay spetial attention to problems with discontinuous solutions or shocks. A fundamental concept in this context is that of upwind discretization, basic to all this family of techniques. Therefore, a detailed description of the first order upwind scheme will be presented as well as the main lines leading to second order extensions. These methods were gradually gaining presence in the context of the shallow water equations years later of their successful performance in the field of gas dynamics [4, 3]. The non-linerity of both inviscid Euler and shallow water equations may complicate their solutions with the appearance of discontinuities reflecting physical phenomena such as shock waves, hydraulic jumps and bores. The numerical technique applied is crucial in these cases and the improvement experienced in both fileds is similar [5, 6]. In general, high resolution techniques are based on mathematical theories well established for homogeneous linear problems or 1D homogeneous nonlinear problems. Among them, Total Variation Diminishing Techniques (TVD) have proved very powerful [7].

Once some of the schemes suitable for the numerical modelling of the homogeneous shallow water equations have been described, the problem of the source terms must be faced. This will be done first for the one-dimensional case and also for the two-dimensional case. The presence of high bottom slopes, important roughness coefficients and strong variations within irregular topography represents a great challenge to the model and can lead to additional numerical errors. The field study devoted to partial differential equations with source terms is becoming more and more active since it is present in an important number of problems of practical interest. Roe [4] pointed out the convenience of discretizing source terms and fluxes in a similar form. In [8, 5] and [9], decomposition methods to solve advection equations with non-linear source terms are proposed. In the line of these authors, a way to discretize the source terms when applying upwind techniques will be presented. If the source terms are discretized in a pointwise manner, numerical oscilations and inaccuracies can arise both in steady and unsteady calculations. Both the flux and the bed slope appear in the shallow water model through their spatial derivatives and that is the reason why their discretization must be analogous in order to achieve perfect balance at least in quiescent steady state. Due to the initial conditions in problems of advance over dry bed and mainly in presence of important roughness, the friction terms become dominant in the equations and can lead to numerical errors if not properly treated.

Numerical models of overland flow have been applied to a number of practical problems of interest in Engineering, including overland hydrology, open channel management and surface irrigation. In the domain of river flow, this type of numerical models are particularly interesting for the simulation of flood waves and their interaction with existing structures. Some of the more advanced approaches, reported for two-dimensional hydraulic problems are based on operator splitting. This consists of regarding the 2D situation as two 1D problems and then using a one-dimensional scheme for each problem separately. Usually, Strang's approximation [10] is followed, and a Cartesian 2D mesh is often required. Other proposed techniques are the use of the method of characteristics in 2D [11], Eulerian-Lagrangian techniques, or alternatively, the finite element technique [12].

In the two-dimensional approach presented in this work, the spatial domain of integration is covered by a set of quadrilateral or triangular cells, not necessarily aligned with the coordinate directions. A discrete approximation to the integral form of the equations is applied in every cell so that the volume integrals represent integrals over the area of the cell and the surface integrals represent the total flux through the cell boundaries.

## 2 1D shallow water equations: Properties

Many hydraulic situations can be described by means of a one-dimensional model, either because a more detailed resolution is unnecessary or because the flow is markedly one-dimensional. The fundamental hypothesis implied in the numerical modelling of river flows are formalized in the equations of unsteady open channel flow. They can be derived, for instance, from mass and momentum control volume analysis and are a simplified model of a very complex phenomenon but they are considered an adequate description for most of the problems associated with open channel and river flow modelling under the St. Venant hypotheses [2]. The 1D unsteady shallow water flow can be written in the form

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = \mathbf{R}, \tag{1}$$

with

$$\begin{aligned} \mathbf{U} &= (A, Q)^T, \\ \mathbf{F} &= (Q, \frac{Q^2}{A} + gI_1)^T, \\ \mathbf{R} &= (0, gI_2 + gA(S_0 - S_f))^T, \end{aligned}$$

which emphasizes the conservative character of the system in the absence of source terms. The effects of the wind as well as those of the Coriolis force have been neglected and no lateral inflow/outflow is considered. In (2.1), $A$ is the wetted cross sectional area, $Q$ is the discharge and $g$ is the acceleration due to gravity. $I_1$ represents a hydrostatic pressure force term as described in [2]

$$I_1 = \int_0^{h(x,t)} (h - \eta) b(x, \eta) d\eta$$

in terms of the surface water level $h(x, t)$ and the breadth

$$b(x, \eta) = \frac{\partial A(x, t)}{\partial \eta}.$$

The pressure forces can have a component in the direction of the main stream due to the reaction of the walls in case of variations in shape along this direction. The amount of this force depends on the cross sectional variation for constant depth. It is important to note that the validity of this approach is linked to the hypothesis of gradual variation. If sudden expansions or contractions take place, the approach is not valid. $I_2$ accounts for the pressure forces in a volume of constant depth $h$ due to longitudinal width variations.

$$I_2 = \int_0^{h(x,t)} (h - \eta) \frac{\partial b(x, \eta)}{\partial x} d\eta.$$

According to both definitions, and following Leibnitz's derivation rule,

$$\frac{\partial I_1}{\partial x} = I_2 + A \frac{\partial h}{\partial x}. \tag{2}$$

The mass force is the projection of the weight of the volume of water in the direction of the stream. The bed slope is the spatial partial derivative of the bottom elevation $z$,

$$S_0 = -\frac{\partial z}{\partial x}.$$

This representation is a consequence of the hypothesis made about the size of the bottom slope. The friction term represents the action of the shear between the fluid and the solid walls. $S_f$ stands for the energy grade line and is defined, for example, in terms of the Manning's roughness coefficient $n$ [13]:

$$S_f = \frac{Q\,|Q|\,n^2}{A^2 R^{\frac{4}{3}}},$$

with $R = A/P$, $P$ being the wetted perimeter. Other forms of $S_f$ can equally well be used.

Other formulations are possible and frequently encountered. A simpler, non-conservative and very common expression of the system, useful in situations of continuous flows, is

$$\begin{aligned} \frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} &= 0, \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x}(\frac{Q^2}{A}) + gA\frac{\partial h}{\partial x} &= gA(S_0 - S_f). \end{aligned} \tag{3}$$

In the particular case of rectangular channels of constant breadth, they reduce to the original equations presented by St. Venant in 1871. Being simpler, the equations in the ideal homogeneous case admit exact solutions that have been traditionally used to test the performance of the numerical techniques. It must be noted that they keep the non linear convective character and, therefore, admit discontinuous (weak) solutions [1].

In those cases in which $\mathbf{F} = \mathbf{F}(\mathbf{U})$ it is possible to rewrite the conservative system in the form

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{J}\frac{\partial \mathbf{U}}{\partial x} = \mathbf{R}, \tag{4}$$

where the Jacobian matrix of the system (4) is

$$\mathbf{J} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{pmatrix} 0 & 1 \\ c^2 - u^2 & 2u \end{pmatrix},$$

where $u = Q/A$ is the cross section averaged water velocity and $c = \sqrt{gA/b}$ is the celerity of the small amplitude surface waves. It is analogous to the speed of sound in gases and contains the essence of the compressibility associated to the deformability of the free surface. At the same time, it is the basis of the definition of the Froude number, $Fr = u/c$, dimensionless number governing this kind of flow, which, also in analogy to the Mach number, allows for a classification in three flow regimes: subcritical ($Fr < 1$), supercritical ($Fr > 1$) and critical ($Fr = 1$).

The system of equations (4) is a hyperbolic system of partial differential equations. Therefore, the Jacobian matrix $\mathbf{J}$ presents interesting properties closely linked to the physics of the problem represented by the mathematical model. The matrix can be made diagonal by means of the set of eigenvalues, which are real and represent the speed of propagation of the information. At the same time, the matrix has a set of linearly independent eigenvectors.

The Jacobian's eigenvalues can be obtained from $|a\mathbf{I} - \mathbf{J}| = 0$ and are

$$a^{1,2} = u \pm c. \tag{5}$$

They represent the speed of propagation of the perturbations and hence are the convective wave velocities. If the Jacobian was a constant matrix, the system would be linear and decoupled. Being a variable matrix in terms of the dependent variables, the system is nonlinear and coupled, and the advection velocities can change of sign and value locally.

The eigenvectors can be obtained from $\mathbf{Je} = a\mathbf{e}$ and are of the form

$$\mathbf{e}^{1,2} = (1, u \pm c)^T. \tag{6}$$

This form of the equations will be particularly useful in the context of upwind schemes. At the same time it is directly related to the theory of characteristics since it enables the diagonalization of the Jacobian. This is achieved by means of the matrix $\mathbf{P}$ made of the column eigenvectors so that

$$\mathbf{J} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}, \qquad \mathbf{\Lambda} = \mathbf{P}^{-1}\mathbf{J}\mathbf{P}. \tag{7}$$

Matrix $\mathbf{P}$ is also the matrix responsible for the change of variables

$$\mathbf{W} = \mathbf{P}^{-1}\mathbf{U}, \tag{8}$$

$$\mathbf{U} = \mathbf{P}\mathbf{W}. \tag{9}$$

So that the system may be rewritten as follows

$$\frac{\partial \mathbf{W}}{\partial t} + \mathbf{\Lambda}\frac{\partial \mathbf{W}}{\partial x} = 0. \tag{10}$$

Due to the construction of matrix $\mathbf{P}$, (9) is equivalent to a linear combination of the Jacobian's eigenvectors in which the coefficients are precisely the characteristic variables

$$\mathbf{U} = \sum_{k=1}^{m} \alpha_k \mathbf{e}_k. \tag{11}$$

In the new variables $\mathbf{W}$, the system is decoupled and formed by a set of scalar equations like

$$\frac{\partial w_k}{\partial t} + a_k \frac{\partial w_k}{\partial x} = 0, \tag{12}$$

as many equations as eigenvalues in the Jacobian matrix $\mathbf{J}$, that is, two in our case. This new formulation is the so called characteristic formulation that belongs to next section.

*2.1 Characteristic formulation*

The characteristic formulation in 1D leads naturally to the method of characteristics. It can be seen as a way to move from a problem governed by partial differential equations to another problem based on ordinary differential equations. The application of this method to unsteady open channel flow can be found in classical references such as [14, 1]. Apart from the derived method of solution, the characteristic formulation is essential to understand the behaviour of the hyperbolic system solutions.

An alternative derivation of the system (12) can be performed starting from the governing equations for a horizontal and frictionless unit width rectangular channel. In that case, the original equations can be manipulated to:

$$\begin{aligned} \frac{\partial h}{\partial t} + u\frac{\partial h}{\partial x} + h\frac{\partial u}{\partial x} &= 0, \\ \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + g\frac{\partial h}{\partial x} &= 0. \end{aligned} \tag{13}$$

Using that $c^2 = gh$, and by means of some simple transformations, Eq. (13) can be rewritten as

$$\begin{aligned} \frac{\partial}{\partial t}(u+2c) + (u+c)\frac{\partial}{\partial x}(u+2c) &= 0, \\ \frac{\partial}{\partial t}(u-2c) + (u-c)\frac{\partial}{\partial x}(u-2c) &= 0. \end{aligned} \tag{14}$$

This result is analogous to Eq. (12). The interpretation of this form of the equations is that the quantity $u + 2c$ is constant along a direction in the $(x, t)$ plane given by the local slope $u + c$ and the same is true for $u - 2c$ along the line of slope $u - c$.

$$\begin{aligned} \frac{dx}{dt} = u + c \quad &\Rightarrow \quad d(u+2c) = 0 \Rightarrow u + 2c = \text{cte}, \\ \frac{dx}{dt} = u - c \quad &\Rightarrow \quad d(u-2c) = 0 \Rightarrow u - 2c = \text{cte}. \end{aligned} \tag{15}$$

The directions in (15) are called characteristic directions ($C^+$ positive, and $C^-$ negative). The quantities conserved along them are the Riemann invariants ($J^+$ and $J^-$, respectively). In a flow domain governed by a system of hyperbolic equations such as the shallow water equations there exist two characteristic directions at every point in the plane $(x, t)$ with the mentioned properties.

In the more general and realistic case of having source terms in the original equations, the invariants become quasi-invariants along the characteristic lines and the form in which they vary is determined by the directional time integral of the source terms. The interaction between the characteristic lines and the spatial boundaries is interesting for the analysis of the required boundary conditions.

## 3 2D mathematical model

It is generally accepted that the unsteady flow of water in a two-dimensional space may be described by the shallow water equations, which represent mass and momentum conservation and can be obtained by depth averaging the Navier-Stokes equations in the vertical direction. This leads to a 2D formulation in terms of depth averaged quantities and the water depth itself [15] and, neglecting diffusion of momentum due to turbulence, they form the following system of equations:

$$\frac{\partial h}{\partial t} + \frac{\partial hu}{\partial x} + \frac{\partial hv}{\partial y} = 0, \tag{16}$$

$$\frac{\partial hu}{\partial t} + \frac{\partial hu^2}{\partial x} + \frac{\partial huv}{\partial y} = fhv + h\tau_{sx} - gh\frac{\partial H}{\partial x} + c_f u\sqrt{u^2+v^2}, \tag{17}$$

$$\frac{\partial hv}{\partial t} + \frac{\partial huv}{\partial x} + \frac{\partial hv^2}{\partial y} = -fhu + h\tau_{sy} - gh\frac{\partial H}{\partial y} + c_f v\sqrt{u^2+v^2}. \tag{18}$$

$f$ represents the Coriolis parameter $f = 2\Omega \sin\phi$ and contributes as a non-inertial volumetric force when geophysical problems of planetary scale are considered. It contains the effect of the Earth rotation on a moving fluid ($\Omega$ being the rotation angular velocity and $\phi$ being the geographic latitude). The relative importance of this term is controlled by the Rosby number. The shear stresses acting on the free surface are due to viscosity and the dynamic boundary condition requires that they are continuous across the surface, that is, their value at the internal part is equal to the external value imposed by the wind. This is the way to include the effect of the wind and is usually modelled using

$$\tau_s = \rho c_W W^2, \tag{19}$$

a semiempirical formula where $W$ is the module of the wind velocity and $c_W$ is a coefficient depending on the wind direction. Both the magnitude and direction of the wind force are determined by the atmospheric flow. The coefficient $c_f$ appearing in the friction term is normally expressed in terms of the Manning $n$ or the Chézy roughness factor

$$c_f u\sqrt{u^2+v^2} = \frac{n^2 u\sqrt{u^2+v^2}}{h^{\frac{4}{3}}} \quad , \quad c_f v\sqrt{u^2+v^2} = \frac{n^2 v\sqrt{u^2+v^2}}{h^{\frac{4}{3}}}. \tag{20}$$

The roughness coefficient $n$ is in principle dependent on the nature of boundary solid surfaces, but also on the flow Reynolds number, although the latter factor is normally neglected.

The terms originated from the depth average of the pressure gradient are $g\partial H/\partial x$, $g\partial H/\partial y$, which, using $H = h + z_b$, can be written as

$$g\frac{\partial H}{\partial x} = g\frac{\partial h}{\partial x} + g\frac{\partial z_b}{\partial x} \ \ , \ qquadg\frac{\partial H}{\partial y} = g\frac{\partial h}{\partial y} + g\frac{\partial z_b}{\partial y}. \tag{21}$$

The bottom level variations are expressed in the form of a slope as

$$S_{0x} = -\frac{\partial z_b}{\partial x}, \qquad S_{0y} = -\frac{\partial z_b}{\partial y}. \tag{22}$$

And the same notation is applied to the friction terms, using the energy grade slopes.

$$S_{fx} = \frac{c_f u\sqrt{u^2+v^2}}{gh}, \qquad S_{fy} = \frac{c_f v\sqrt{u^2+v^2}}{gh}. \tag{23}$$

Finally the following form can be written for the system of shallow water equations

$$\frac{\partial h}{\partial t} + \frac{\partial hu}{\partial x} + \frac{\partial hv}{\partial y} = 0, \tag{24}$$

$$\frac{\partial hu}{\partial t} + \frac{\partial hu^2}{\partial x} + gh\frac{\partial h}{\partial x} + \frac{\partial huv}{\partial y} = gh(S_{0x} - S_{fx}), \tag{25}$$

$$\frac{\partial hv}{\partial t} + \frac{\partial huv}{\partial x} + \frac{\partial hv^2}{\partial y} + gh\frac{\partial h}{\partial y} = gh(S_{0y} - S_{fy}). \tag{26}$$

The system can be rewritten in conservative form, that is, in the closest form to a system of conservation laws as

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla\left(\mathbf{F}, \mathbf{G}\right) = \mathbf{R}, \tag{27}$$

where

$$\mathbf{U} = \begin{pmatrix} h \\ hu \\ hv \end{pmatrix}, \qquad \mathbf{F} = \begin{pmatrix} hu \\ hu^2 + g\frac{h^2}{2} \\ huv \end{pmatrix},$$
$$\mathbf{G} = \begin{pmatrix} hv \\ huv \\ hv^2 + g\frac{h^2}{2} \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 0 \\ gh\left(S_{0x} - S_{fx}\right) \\ gh\left(S_{0y} - S_{fy}\right) \end{pmatrix}. \tag{28}$$

$\mathbf{U}$ represents the vector of conserved variables ($h$ water depth, $hu$ and $hv$ unit discharges along the coordinate directions $x$, $y$ respectively), $\mathbf{F}$ and $\mathbf{G}$ are the fluxes of the conserved variables across the edges of a control volume and are formed by the convective flux and the hydrostatic pressure gradient. $\mathbf{R}$, contains the sources and sinks of momentum along the two coordinate directions.

Hence, (27) is a system of coupled and nonlinear partial differential equations. It can also be written in a quasi conservative form

$$\frac{\partial \mathbf{U}}{\partial t} + (\mathbf{A}, \mathbf{B})\nabla\mathbf{U} = \mathbf{R}. \tag{29}$$

The Jacobian matrices are

$$\mathbf{A} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} = \begin{pmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix}, \quad \mathbf{B} = \frac{\partial \mathbf{G}}{\partial \mathbf{U}} = \begin{pmatrix} 0 & 0 & 1 \\ -uv & v & u \\ c^2 - v^2 & 0 & 2v \end{pmatrix}. \tag{30}$$

The non conservative form (29) of the two dimensional system of equations is less useful than the analogous in one dimension from the point of view of the numerical resolution because a simultaneous diagonalisation of **A** and **B** is not possible. This is the main reason why most schemes relay on a normal flux formulation. A characteristic form of the equations can also be found in two dimensions. In the 1D case, the Jacobian diagonalization led to the direct identification of characteristic directions, characteristic variables and their governing equations. The existence and properties of the two different Jacobians in 2D only allows a partial diagonalization of the system. A set of special directions and surfaces in the space $(x, y, t)$ can be identified along which differential expressions derived from the original system hold. These are called compatibility relations.

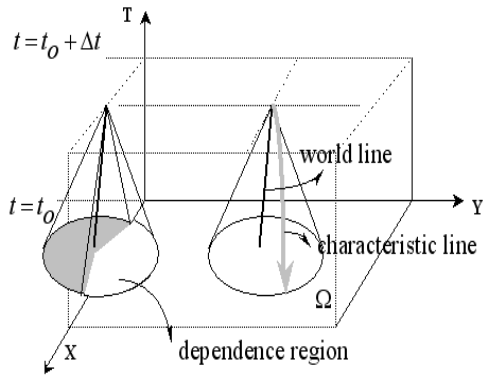


Figure 1.— Characteristic cones.

When dealing with the two-dimensional set of shallow water equations, the characteristic surface cone 1, is centred on the so called world line and generated by an infinite set of tangent bicharacteristic lines that can be expressed in terms of the polar angle and are trajectories carrying information forward in time. When the 1D shallow water equations are considered, two variables, depth, and velocity for instance, are unknown at every point, and two characteristic curves, in the $(x, t)$ plane are used to find the solution. In the 2D case there are three unknowns, water depth and two velocity components, and therefore three equations are required. The information used to find the solution in first order of approximation travels through the world line and two selected bicharacteristic curves. In the first order upwind scheme, the eigenvalues play a role similar to that of the bicharacteristic curves, and that of the world line but computed at every cell edge and contributing to the discretization of the three conservation equations (mass and momentum). Depending on the dimensionless normal Froude number, the characteristic cone has a different shape. At a given cell, a characteristic cone can be defined and the value of the three unknowns at the cell can be updated using information carried by the ingoing characteristic lines. It is important to insist that both the bicharacteristic lines and the negative eigenvalues carry useful information for the updating at a cell only when their influence can be computed from known values of the variables within the domain at the old time level.

It is frequent to define the tensorial flux $\mathbf{E} = (\mathbf{F}, \mathbf{G})^T$ in order to introduce the integral form of the equations over a fixed volume $\Omega$,

$$\frac{\partial}{\partial t}\int_\Omega \mathbf{U}\ d\Omega + \int_\Omega (\nabla \mathbf{E})\ d\Omega = \int_\Omega \mathbf{R}\ d\Omega. \tag{31}$$

This form of the equations of motion is more general and anticipates the finite volume technique of discretization that will be applied. The application of Gauss's theorem to the second integral of the left hand side of (2.16) allows us to rewrite it as

$$\frac{\partial}{\partial t}\int_\Omega \mathbf{U}\ d\Omega + \oint_S (\mathbf{E}\cdot\mathbf{n})\ ds = \int_\Omega \mathbf{R}\ d\Omega, \tag{32}$$

where $S$ denotes the surface surrounding the volume $\Omega$ and $\mathbf{n}$ is the unit outward normal vector.

The Jacobian matrix, $\mathbf{J}_n$, of the normal flux $(\mathbf{E}\cdot\mathbf{n})$ present in (2.17) is evaluated as

$$\mathbf{J}_n = \frac{\partial \mathbf{E}\cdot\mathbf{n}}{\partial \mathbf{U}} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} n_x + \frac{\partial \mathbf{G}}{\partial \mathbf{U}} n_y,$$

and can be expressed as

$$\mathbf{J_n} = \begin{pmatrix} 0 & n_x & n_y \\ (gh - \frac{q_x^2}{h^2})n_x - \frac{q_x q_y}{h^2} n_y & \frac{q_y}{h} n_y + \frac{2q_x}{h} n_x & \frac{q_x}{h} n_y \\ (gh - \frac{q_y^2}{h^2})n_y - \frac{q_x q_y}{h^2} n_x & \frac{q_y}{h} n_x & \frac{q_x}{h} n_x + \frac{2q_y}{h} n_y \end{pmatrix}.$$

where $q_x = hu$ and$q_y = hv$. The eigenvalues of $\mathbf{J}_n$ are a representation of the characteristic speeds $a^l$.

$$\begin{aligned} a^1 &= un_x + vn_y + c, \\ a^2 &= un_x + vn_y, \\ a^3 &= un_x + vn_y - c. \end{aligned}$$

The corresponding eigenvectors are

$$\mathbf{e}^1 = \begin{pmatrix} 1 \\ u + cn_x \\ v + cn_y \end{pmatrix}, \quad \mathbf{e}^2 = \begin{pmatrix} 0 \\ -cn_y \\ cn_x \end{pmatrix}, \quad \mathbf{e}^3 = \begin{pmatrix} 1 \\ u - cn_x \\ v - cn_y \end{pmatrix}.$$

Two matrices $\mathbf{P}$ and $\mathbf{P}^{-1}$ can be constructed with the property that they diagonalise the Jacobian $\mathbf{J}_n$. From their eigenvectors:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 1 \\ u + cn_x & -cn_y & u - cn_x \\ v + cn_y & cn_x & v - cn_y \end{pmatrix}, \quad \mathbf{P}^{-1} = \frac{1}{2c}\begin{pmatrix} c + \mathbf{u}\cdot\mathbf{n} & n_x & n_y \\ 2(un_y - vn_x) & -2n_y & 2n_x \\ c + \mathbf{u}\cdot\mathbf{n} & -n_x & -n_y \end{pmatrix},$$

$$\mathbf{J}_n = \mathbf{P\Lambda P}^{-1},$$

where $\mathbf{\Lambda}$ is a diagonal matrix with eigenvalues in the main diagonal.

Denoting by $\mathbf{U}_i$ and $\mathbf{R}_i$ the average value of the flow variables and source terms respectively over the volume $i$ at a given time, from (2.17) the following conservation equation can be written for every cell:

$$\frac{\partial \mathbf{U}_i}{\partial t}\ \Omega_i + \oint_{S_i} (\mathbf{E}\cdot\mathbf{n})\ ds = \mathbf{R}_i\ \Omega_i. \tag{33}$$

The finite volume procedure defined above is completely general. A mesh fixed in time is assumed and the contour integral is approached via a mid-point rule, i.e., a numerical flux is defined at the mid-point of each edge, giving

$$\oint_{S_i} (\mathbf{E}\cdot\mathbf{n})\ ds \approx \sum_{k=1}^{NE} (\mathbf{E}^*_{wk}\cdot\mathbf{n}_{wk})\ dC_{wk}, \tag{34}$$

where $w_k$ represents the index of edge $k$ of the cell, $NE$ is the total number of edges in the cell ($NE = 3$ for triangles, $NE = 4$ for quadrilaterals). The vector $\mathbf{n}_{w_k}$ is the unit outward normal, $dC_{w_k}$ is the length of the side, and $(\mathbf{F},\mathbf{G})^*_{w_k}$ is the numerical flux tensor.

Different implementations arise depending on the numerical scheme used and, consequently, on the numerical flux $\mathbf{E}^*$.

In the next section, this approach will be particularized to the 1st order Roe's scheme and to MacCormack's scheme where first, the numerical flux corresponding to the classical method will be described and then the TVD modification derived.

## 4 1D numerical techniques

Two forms of describing flow motion are well kown in Fluid Mechanics: Lagrangian and Eulerian. What is essentially done in the Lagrangian form is to follow the particles individual history. In the Eulerian description, the interest is not focused on particle but on fixed points in space. In a totally parallel form, it is possible to distinguish two families of numerical techniques for the resolution of the differential equations governing fluid motion. Lagrangian techniques are those dealing with the equations developed under this philosophy. The discretization used relays on a finite number of particles of known position at the initial time. On the other hand, Eulerian techniques furnish information of the evolution of the variables on a discrete and fixed set of spatial points called grid. The methods included here belong to this category.

There exists a family of techniques for hyperbolic equations which cannot be totally labelled as Lagrangian or Eulerian: those based on the existence of characteristic directions.

These directions have properties that, in some cases, form the basis of an alternative resolution method. This kind of methods are often called Semi-lagrangian since, despite using a Eulerian grid, they manage directional information. Different Semi-lagrangian methods arise depending on the technique used for integrating the trajectory of the characteristic curve or the interpolation method applied. The classical method of characteristics does also fall in this category.

Among Eulerian schemes, a first division can be made according to the kind of time integration: explicit and implicit methods. If the spatial discretization is used as a criterion, two main groups can be made: central and upwind schemes. Independently of these categories, there is a very important class of methods, that of conservative methods. If there is an interest to deal with unsteady problems that may contain rapidly varied flows with possible presence of transcritical flow or formation of surges, attention must be focused on techniques able to progress automatically and correctly towards a weak or discontinuous solution of the equations [16]. They have the important property of ensuring a correct approximation to a conservative equation or system of equations and are hence well adapted to flow simulations with discontinuities. In general, conservative schemes can be cast under the following discrete form

$$\frac{\Delta \mathbf{U}_i}{\Delta t} = \mathbf{R}_i^* - \frac{1}{\Delta x_i}\left(\mathbf{F}_{i+1/2}^* - \mathbf{F}_{i-1/2}^*\right),$$

in which $\mathbf{R}^*$ and $\mathbf{F}^*$ are respectively the numerical source and the numerical flux to be defined in every case. The schemes used for the numerical computations in this work fall mainly into the category of conservative explicit methods.

The linear study of the numerical stability produces the following condition on the time step size of the explicit schemes in 1D

$$CFL = \frac{\Delta t}{\Delta x}\max(a_k) \leq 1, \tag{35}$$

where $a_k$ are the eigenvalues of matrix $\mathbf{J}$. This condition is directly related to the existence of well defined domain of dependence and region of influence in hyperbolic problems. There is a long list of finite difference techniques suitable for the numerical solution of the equations presented. Some representative techniques will be briefly presented here

### *4.1 Lax-Friedrichs scheme*

This is a first order in space and time central finite difference technique. For a general homogeneous conservation system

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0, \tag{36}$$

the procedure to update one time step $\Delta t$ the interior points $2, \ldots, N-1$ of a regular grid is based on a nodal updating:

$$\mathbf{U}_i^{n+1} = \alpha \mathbf{U}_i^n + \frac{1-\alpha}{2}(\mathbf{U}_{i+1}^n + \mathbf{U}_{i-1}^n) - \frac{\Delta t}{2\Delta x}(\mathbf{F}_{i+1}^n - \mathbf{F}_{i-1}^n), \qquad 0 \leq \alpha < 1.$$

The value $\alpha = 1$ renders the scheme unstable. More stability and more numerical diffusion are introduced as $\alpha$ approaches zero. A value $\alpha = 0.1$ is usually adopted. It is easy to implement and very robust. It is a conservative scheme that admits the following numerical flux

$$\mathbf{F}_{i+1/2}^* = \frac{1}{2}\left(\mathbf{F}_{i+1} + \mathbf{F}_i - (1-\alpha)\frac{\Delta x_i}{\Delta t}(\mathbf{U}_{i+1} - \mathbf{U}_i)\right).$$

This method has been extensively studied and applied to the shallow water equations in [17].

### *4.2 First order Roe's Scheme*

Upwind schemes are based on the idea of discretizing the spatial derivatives so that information is taken from the side it comes. Hence, a sense of propagation is implied and these techniques are well adapted to advection dominated problems. The extension of the technique to a non-linear system like

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = 0,$$

exploits the form

$$\frac{\partial \mathbf{U}}{\partial t} + \mathbf{J}\frac{\partial \mathbf{U}}{\partial x} = 0,$$

and the property of the Jacobian that

$$\mathbf{J} = \mathbf{P\Lambda P}^{-1},$$

where $\mathbf{\Lambda}$ is the diagonal matrix with the eigenvalues in the main diagonal and $\mathbf{P}$ is the matrix made of the right eigenvectors. If the system was linear, it would be possible to decouple it in a straightforward manner and apply the scalar technique to every equation as follows:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x_i}\left((\mathbf{P\Lambda^- P}^{-1})_{i+\frac{1}{2}}^n(\mathbf{U}_{i+1}^n - \mathbf{U}_i^n) + (\mathbf{P\Lambda^+ P}^{-1})_{i-\frac{1}{2}}^n(\mathbf{U}_i^n - \mathbf{U}_{i-1}^n)\right),$$

where

$$\mathbf{\Lambda}^{\pm} = \frac{1}{2}(\mathbf{\Lambda} \pm |\mathbf{\Lambda}|), \qquad \mathbf{J}^{\pm} = \mathbf{P\Lambda^{\pm}P}^{-1}. \tag{37}$$

In the non-linear case, a local linearization is performed according to Roe's approximate Riemann solver [**?**]. At every cell limited by nodes $L$ and $R$, an approximate Jacobian $(\tilde{\mathbf{J}}_{RL})$ is defined satisfying the following conditions:

I) $\tilde{\mathbf{J}}_{RL}$ depends only on the $\mathbf{U}_R$ and $\mathbf{U}_L$ states,

II) $(\mathbf{F}_R - \mathbf{F}_L) = \tilde{\mathbf{J}}_{RL}\,(\mathbf{U}_R - \mathbf{U}_L)$,

III) $\tilde{\mathbf{J}}_{RL}$ has real and distinct eigenvalues and a complete set of eigenvectors,

IV) $\tilde{\mathbf{J}}_{RL} = \mathbf{J}\,(\mathbf{U}_R) = \mathbf{J}\,(\mathbf{U}_L)$ if $\mathbf{U}_R = \mathbf{U}_L$.

The expressions for the eigenvalues $\tilde{a}^k$ and eigenvectors $\tilde{\mathbf{e}}^k$ are similar to those of the original Jacobian but evaluated at average values in $(L, R)$.

$$\tilde{a}^{1,2} = \tilde{u} \pm \tilde{c}, \qquad \tilde{\mathbf{e}}^{1,2} = (1, \tilde{u} \pm \tilde{c})^T,$$

where the average velocity and celerity at the interface $(L, R)$ are

$$\tilde{u} = \frac{Q_R\sqrt{A_L} + Q_L\sqrt{A_R}}{\sqrt{A_L A_R}(\sqrt{A_R} + \sqrt{A_L})}, \qquad \tilde{c} = \frac{1}{2}(c_L + c_R).$$

The basic idea is to calculate $\delta\mathbf{U}$ at every interface and propagate the different $k$ waves according to the sign of their celerities (eigenvalues) and the values of the local $CFL$ numbers. The scheme is conservative and the system version of the numerical flux is

$$\mathbf{F}^*_{i+\frac{1}{2}} = \frac{1}{2}(\mathbf{F}_{i+1} + \mathbf{F}_i) - \frac{1}{2}\left((\tilde{\mathbf{P}}\left|\tilde{\mathbf{\Lambda}}\right|\tilde{\mathbf{P}}^{-1})_{i+1/2}(\mathbf{U}_{i+1} - \mathbf{U}_i)\right). \tag{38}$$

This method has been studied and applied to the shallow water equations in [19], [20] and [18] for instance.

### *4.3 Lax-Wendroff scheme*

In a search for stable and more accurate shock capturing numerical schemes, P. Lax and B. Wendroff [21] proposed the idea of combining the spatial and temporal discretization in order to globally achieve second order. Lax-Wendroff's scheme is an explicit second order method. It was further simplified by some authors in order to avoid evaluation of the celerity (Jacobian) at an intermediate position. In the case of homogeneous systems of equations, the same guidelines lead to the following scheme

$$\begin{aligned} \mathbf{U}_i^{n+1} &= \mathbf{U}_i^n - \frac{\Delta t}{2\Delta x}\left(\mathbf{F}_{i+1}^n - \mathbf{F}_{i-1}^n\right) + \\ &+ \frac{(\Delta t)^2}{2(\Delta x)^2}\left[\mathbf{J}_{i+\frac{1}{2}}^n\left(\mathbf{F}_{i+1}^n - \mathbf{F}_i^n\right) - \mathbf{J}_{i-\frac{1}{2}}^n\left(\mathbf{F}_i^n - \mathbf{F}_{i-1}^n\right)\right], \end{aligned} \tag{39}$$

where now $\mathbf{J} = \partial\mathbf{F}/\partial\mathbf{U}$ is the Jacobian matrix of the system. If the system is linear, the matrix is constant $\mathbf{F} = \mathbf{J}\mathbf{U}$ with $\mathbf{J} =$ constant, and, in the non-linear case, $\mathbf{J}\,(\mathbf{U})$ must be evaluated at an intermediate position $\mathbf{J}_{i+\frac{1}{2}} = \mathbf{J}\left(\mathbf{U}_{i+\frac{1}{2}}\right)$. The numerical flux can be written as

$$\mathbf{F}^*_{i+\frac{1}{2}} = \frac{1}{2}\left(\mathbf{F}_{i+1}^n + \mathbf{F}_i^n\right) + \frac{\Delta t}{2\Delta x}\left(\mathbf{J}_{i+\frac{1}{2}}^n\right)^2\left(\mathbf{U}_{i+1}^n - \mathbf{U}_i^n\right). \tag{40}$$

The scheme is non-dissipative for $\mathbf{J}$ = constant, and displays oscillations near strong gradients (shocks). It can also lead to numerical difficulties near critical or sonic points. Several authors have recommended the addition of extra dissipative terms (pseudoviscosity) in these cases, see [17] for instance. Lax-Wendroff's scheme is one of the most frequently encountered in the literature related to classical shock-capturing schemes. Difficulties have been reported when trying to include source terms in the discretization and to keep second order of accuracy at the same time. This was studied in [18].

### *4.4 Second order upwind scheme*

In order to improve the accuracy of the first order upwind scheme, second order can be found by means of a Taylor series as seen in the Lax-Wendroff case. Discretizing now the spatial derivatives in a non-central or upwind way to second order, the following is obtained for the non-linear systems of equations:

$$\begin{aligned}\mathbf{F}^*_{i+1/2} = \underbrace{\mathbf{F}^+_i + \mathbf{F}^-_{i+1}}_{\text{first order}} \; &+ \; \frac{1}{2}\left(1 - \frac{\Delta t}{\Delta x}\tilde{\mathbf{J}}^+_{i-1/2}\right)\left(\mathbf{F}^+_i - \mathbf{F}^+_{i-1}\right) - \\ &- \; \frac{1}{2}\left(1 + \frac{\Delta t}{\Delta x}\tilde{\mathbf{J}}^-_{i+3/2}\right)\left(\mathbf{F}^-_{i+2} - \mathbf{F}^-_{i+1}\right), \end{aligned} \tag{41}$$

where $\mathbf{F}^{\pm}$ and $\mathbf{A}^{\pm}$ are defined like in (3.13).

### *4.5 High resolution TVD schemes*

There is a recent group of methods combining the advantages of both the second order accuracy and the first order smoothness. They will be briefly presented in what follows. The fundamental concept to Total Variation Diminishing (TVD) methods [8], [7] is the definition of the Total Variation of a continuous function $u(x,t)$.

$$TV\left(u(x,t)\right) = \int |du| = \int \left|\frac{\partial u}{\partial x}\right| dx, \tag{42}$$

which has the property of not increasing in time (Harten 1984), i.e.,

$$TV\left(u(x,t_2)\right) \leq TV\left(u(x,t_1)\right), \quad \text{with} \quad t_2 > t_1. \tag{43}$$

Any numerical method designed to have this property at the discrete level will remain free of oscillations and new extrema no matter the accuracy achieved. The total variation of a discrete function is expressed as

$$TV\left(u^n\right) = \sum_i \left|\delta u^n_{i+1/2}\right| \tag{44}$$

and a numerical approximation will be TVD whenever

$$TV\left(u^{n+1}\right) \leq TV\left(u^n\right). \tag{45}$$

Among the schemes presented so far the first order schemes are already TVD when working within the stability limits (Hirsch, 1990). The second order schemes can be made TVD by means of a technique that combines first order parts and second order corrections. The second order upwind scheme, for instance, can be re-expressed as follows

$$\begin{aligned} u_i^{n+1} &= u_i^n - \frac{\Delta t}{\Delta x}\left[\delta f_{i-\frac{1}{2}}^+ + \delta f_{i+\frac{1}{2}}^-\right] + \\ &+ \frac{\Delta t}{2\Delta x}\varphi\left(r_{i+\frac{3}{2}}^-\right)\left[1 - \frac{\Delta t}{\Delta x}a_{i+\frac{3}{2}}^-\right]\delta f_{i+\frac{3}{2}}^- - \frac{\Delta t}{2\Delta x}\varphi\left(r_{i+\frac{1}{2}}^-\right)\left[1 - \frac{\Delta t}{\Delta x}a_{i+\frac{1}{2}}^-\right]\delta f_{i+\frac{1}{2}}^- - \\ &- \frac{\Delta t}{2\Delta x}\varphi\left(r_{i-\frac{1}{2}}^+\right)\left[1 + \frac{\Delta t}{\Delta x}a_{i-\frac{1}{2}}^+\right]\delta f_{i-\frac{1}{2}}^+ + \frac{\Delta t}{2\Delta x}\varphi\left(r_{i-\frac{3}{2}}^+\right)\left[1 + \frac{\Delta t}{\Delta x}a_{i-\frac{3}{2}}^+\right]\delta f_{i-\frac{3}{2}}^+ \end{aligned}$$

The factors affecting the second order correction terms are called flux limiters $\varphi(r)$. They are non-linear functions of the local gradients of the solution working so that the methods becomes of first order in the vicinity of strong gradients but remains of second order in regions of smooth flow.

To achieve this goal, the gradient at a cell interface is compared to the gradient at the neighbouring interface given by the sign of the advection speed. The argument of the limiter is:

$$r_{i+\frac{1}{2}}^- = \frac{\left[\left(1 + \frac{\Delta t}{\Delta x}a^-\right)\delta f^-\right]_{i-\frac{1}{2}}}{\left[\left(1 + \frac{\Delta t}{\Delta x}a^-\right)\delta f^-\right]_{i+\frac{1}{2}}}, \qquad r_{i-\frac{1}{2}}^+ = \frac{\left[\left(1 - \frac{\Delta t}{\Delta x}a^+\right)\delta f^+\right]_{i+\frac{1}{2}}}{\left[\left(1 - \frac{\Delta t}{\Delta x}a^+\right)\delta f^+\right]_{i-\frac{1}{2}}}. \tag{46}$$

Then for instance, if the limiter is nil, $\varphi(r) = 0$ there are no second order corrections in the scheme and the first order upwind is recovered. When the limiter takes the value $\varphi(r) = 1$, the method becomes the second order upwind scheme, and for $\varphi(r) = r$, the method becomes the Lax-Wendroff scheme. For any other value between these limits [22] the method behaves with the required properties. To guarantee them, the limiter functions must have certain properties. Some of the most widespread are given below:

- Van Leer: $\varphi(r) = \dfrac{r + |r|}{1 + r}$
- Minmod: $\varphi(r) = \max\left[0, \min(r, 1), \min(1, r)\right]$
- Superbee: $\varphi(r) = \max\left[0, \min(2r, 1), \min(2, r)\right]$
- Van Albada: $\varphi(r) = \dfrac{2r}{1 + r^2}$

The application of these techniques to 1D shallow water problems with source terms was reported by [18].

## 5 2D numerical techniques

We shall concentrate on the first order upwind method. An important feature of the 1D upwind schemes for non-linear systems of equations is exploited here. This is the definition of the approximated flux jacobian, $\tilde{\mathbf{J}}$, constructed at the edges of the cells. The 1D philosophy is followed along the normal direction to the cell walls, making use of the normal numerical fluxes. Once this matrix has been defined, the numerical flux across each edge $w_k$ of the computational cells $L$ on the left and $R$ on the right of a cell in a 2D domain is

$$(\mathbf{F},\mathbf{G})^* \cdot \mathbf{n} = \frac{1}{2}\left[(\mathbf{F},\mathbf{G})_R \cdot \mathbf{n} + (\mathbf{F},\mathbf{G})_L \cdot \mathbf{n} - |\tilde{\mathbf{J}}_{RL}|(\mathbf{U}_R - \mathbf{U}_L)\right]. \tag{47}$$

Here $(\tilde{\mathbf{J}}_{RL})$ represents the approximate Jacobian of the normal flux. Note that subscript $k$ will be omitted for the sake of clarity and the following discussion is referred to the cell side $k$.

As suggested by Roe [4], [23] the matrix $\tilde{\mathbf{J}}_{RL}$ has the same shape as $\mathbf{J_n}$ but is evaluated at an average state given by the quantities $\tilde{\mathbf{u}} = (\tilde{u}, \tilde{v})$ and $\tilde{c}$ which must be calculated according to the matrix properties [24]:

1. $\tilde{\mathbf{J}}_{RL} = \tilde{\mathbf{J}}_{RL}(\mathbf{U}_R, \mathbf{U}_L)$.
2. $\mathbf{F}_R - \mathbf{F}_L = \tilde{\mathbf{J}}_{RL}(\mathbf{U}_R - \mathbf{U}_L)$.
3. $\tilde{\mathbf{J}}_{RL}$ has a complete set of real and different eigenvalues and eigenvectors.
4. $\tilde{\mathbf{J}}_{RL}(\mathbf{U}_L, \mathbf{U}_L) = \mathbf{J_n}(\mathbf{U}_L)$ .

The approximate Jacobian matrix is not directly used in the actual method. Instead, the difference in the vector $\mathbf{U}$ across the grid edge is decomposed on the matrix eigenvectors basis as

$$\Delta\mathbf{U} = \mathbf{U}_R - \mathbf{U}_L = \sum_{m=1}^{3} \alpha^m \tilde{\mathbf{e}}^m, \tag{48}$$

where the expression of coefficients $\alpha^m$ are:

$$\alpha^{1,3} = \frac{h_R - h_L}{2} \pm \frac{1}{2\tilde{c}}\left[\left((hu)_R - (hu)_L\right) n_x + \left((hv)_R - (hv)_L\right) n_y - \left(\tilde{u}n_x + \tilde{v}n_y\right)\left(h_R - h_L\right)\right],$$
$$\alpha^2 = \frac{1}{\tilde{c}}\left[\left((hv)_R - (hv)_L - \tilde{v}\left(h_R - h_L\right) n_x\right) - \left((hv)_R - (hv)_L - \tilde{u}\left(h_R - h_L\right) n_y\right)\right]. \tag{49}$$

Matrix $|\tilde{\mathbf{J}}_{RL}|$ is replaced by its eigenvalues and eigenvectors in the product $|\tilde{\mathbf{J}}_{RL}|\left(\mathbf{U}_R - \mathbf{U}_L\right)$ in the form

$$|\tilde{\mathbf{J}}_{RL}|\left(\mathbf{U}_R - \mathbf{U}_L\right) = \sum_{m=1}^{3} |\tilde{a}^m| \alpha^m \tilde{\mathbf{e}}^m. \tag{50}$$

From the eigenvalues of $\mathbf{J}$, those of $\tilde{\mathbf{J}}_{RL}$ have the same form, all in terms of average velocities and celerity. Enforcing the second condition of the matrix $\tilde{\mathbf{J}}_{RL}$ the following expressions for $\tilde{u}$, $\tilde{v}$ and $\tilde{c}$ can be obtained

$$\tilde{u} = \frac{\sqrt{h_R}u_R + \sqrt{h_L}u_L}{\sqrt{h_R} + \sqrt{h_L}}, \quad \tilde{v} = \frac{\sqrt{h_R}v_R + \sqrt{h_L}v_L}{\sqrt{h_R} + \sqrt{h_L}}, \quad \tilde{c} = \sqrt{\frac{g}{2}\left(h_R + h_L\right)}. \tag{51}$$

It has to be stressed at this point that in case of an advancing front over dry bed the average velocities are calculated in the form

$$\tilde{u} = \frac{u_R + u_L}{2}, \quad \tilde{v} = \frac{v_R + v_L}{2}, \tag{52}$$

since the velocity values at the right or left cell are zero. This election is the proposed by the $Q$-scheme of Van Leer [25] for this situation.

The numerical flux normal to each edge of the computational cells becomes

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{A_i}\left(\sum_{k=1}^{NE} \mathbf{E}_k^* \cdot \mathbf{n}_k ds_k\right)_i^n + \Delta t \int_\Omega \mathbf{R}\ d\Omega. \tag{53}$$

This form of updating the variables via a numerical interface flux is common in finite volume methods. It is less common, but also possible, to re-express (53) in a different form by realizing that

$$\Delta\left(\mathbf{E}\cdot\mathbf{n}\right) = \tilde{\mathbf{J}}_{RL}\Delta\mathbf{U} = \tilde{\mathbf{P}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{P}}^{-1}\Delta\mathbf{U} = \tilde{\mathbf{P}}(\tilde{\mathbf{\Lambda}}^{+} + \tilde{\mathbf{\Lambda}}^{-})\tilde{\mathbf{P}}^{-1}\Delta\mathbf{U}, \tag{54}$$

where $\mathbf{\Lambda}^{\pm} = \left(\mathbf{\Lambda} \pm |\mathbf{\Lambda}|\right)/2$, and the previous decomposition represents the splitting of the gradient in left and right traveling parts. For the updating of a single cell, only the in-going contributions are taken into account so that the contour integral of the numerical normal flux is equivalent to the sum of these waves.

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{A_i}\left(\sum_{k=1}^{NE} (\tilde{\mathbf{P}}\tilde{\mathbf{\Lambda}}^{-}\tilde{\mathbf{P}}^{-1}\Delta\mathbf{U})_k ds_k\right)_i^n + \Delta t \mathbf{R}^{*n}_i. \tag{55}$$

For the numerical source, $\mathbf{R}^*$, an approach of the integral of the source term $\mathbf{R}$ over the cell has to be defined. First, it should be noted that the source term vector can be decomposed in two different parts that will be treated separately: the bottom variations $\mathbf{R}^1$ and the friction term $\mathbf{R}^2$, $\mathbf{R} = \mathbf{R}^1 + \mathbf{R}^2$ corresponding to

$$\mathbf{R}^1 = \begin{pmatrix} 0 \\ ghS_{0x} \\ ghS_{0y} \end{pmatrix}, \ \mathbf{R}^2 = \begin{pmatrix} 0 \\ -ghS_{fx} \\ -ghS_{fy} \end{pmatrix}. \tag{56}$$

The first term $\mathbf{R}^1$ accounts for the bed slopes and is the only one containing spatial derivatives. For this reason the discretization procedure will follow the flux term discretization as close as possible as suggested by [25]. The second term $\mathbf{R}^2$ accounts for the friction. Extra terms could be added to take into account infiltration processes, for instance.

An upwind approach has been adopted to model the bottom variations in order to ensure the best balance with the flux terms at least in steady cases. This procedure is analogous in 1D. The flux discretization in Eq. (55) can be used in the same way for the bottom slope terms because both contemplate the same spatial derivative.

For every cell-edge $k$ of cell $\Omega_i$ the discrete source term is decomposed into inward and outward contributions

$$\tilde{\mathbf{R}}_k^1 = \tilde{\mathbf{R}}_k^{1+} + \tilde{\mathbf{R}}_k^{1-},$$

with

$$\tilde{\mathbf{R}}_k^{1\pm} = \tilde{\mathbf{P}} \left( \mathbf{I} \pm |\tilde{\mathbf{\Lambda}}| \tilde{\mathbf{\Lambda}}^{-1} \right) \tilde{\mathbf{P}}^{-1} \tilde{\mathbf{R}}_k^1 = \sum_{m=1}^{3} \beta^{m\pm} \tilde{\mathbf{e}}^m. \tag{57}$$

The average value $\tilde{\mathbf{R}}_k^1$ is computed with

$$\tilde{\mathbf{R}}_k^1 = \begin{pmatrix} 0 \\ g\tilde{h}\Delta z_{bx} \\ g\tilde{h}\Delta z_{by} \end{pmatrix}_k , \tag{58}$$

where $\tilde{h}$ consists of the average obtained from the depth values stored in the left and right cell that share the same edge in each computational cell:

$$\tilde{h} = \frac{1}{2} \left( h_R + h_L \right). \tag{59}$$

and the bed increments in each direction are computed in the form

$$\Delta z_{bx} = -(z_{bR} - z_{bL}) n_x, \quad \Delta z_{by} = -(z_{bR} - z_{bL}) n_y, \quad \mathbf{n} = (n_x, n_y). \tag{60}$$

For every cell $\Omega_i$ the total contribution of the source terms is made of the sum of the parts associated to inward normal velocity at every edge $k$

$$\mathbf{R}_i^{1*} = \sum_{k=1}^{NE} \tilde{\mathbf{R}}_k^{1-}.$$

For that reason we define the numerical source term at cell-edge $k$ as

$$\mathbf{R}_k^{1*} = \tilde{\mathbf{R}}_k^{1-}.$$

The expressions for the $\beta^-$ coefficients are

$$\begin{aligned} \beta^{1-,3-} &= \pm \frac{1}{2\tilde{c}} \left( 1 - \frac{|a_{1,3}|}{a_{1,3}} \right) \left[ \mathbf{R}_2^1 n_x + \mathbf{R}_3^1 n_y \right], \\ \beta^{2-} &= \frac{1}{\tilde{c}} \left( 1 - \frac{|a_2|}{a_2} \right) \left( -\mathbf{R}_2^1 n_y + \mathbf{R}_3^1 n_x \right). \end{aligned} \tag{61}$$

The average value, Eq. (58), proposed in [25], ensures a conservative discretization of this source term.

The numerical scheme for this part is formulated as

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{S_i} \left( \sum_{k=1}^{NE} \left( (\mathbf{F}, \mathbf{G})_k^* \cdot \mathbf{n}_k dC_k - \mathbf{R}_k^{1*} \right) \right)^n. \tag{62}$$

In practical two-dimensional river flow aplications, water is not confined into a vessel but has the freedom to overflow the main channel depending on the flood event. Due to the small initial values (zero actually) for the variables $h$, $u$, and $v$ and the high value of the Manning coefficient required to take into account the effects of the vegetation in the dry field, the friction terms may become dominant and the numerical solution is affected. The main consequence is in the form of numerical stability restrictions different from the well known CFL conditions. This is further complicated by the mathematical difficulties linked to the advance of the front over dry bed.

To avoid this problem, one option is to discretize the friction term in an upwind manner even though it does not show a gradient form. This option does not prove useful in relieving numerical troubles. Back to the pointwise option, two possibilities to avoid this problem can be considered; the first is to ignore the friction terms when the depth of water is smaller than a threshold; another approach is to treat the source terms in a semi-implicit form. In the approximation presented here, the friction terms are calculated in the following way

$$S_{fx} = -(1-\theta)(ghS_{fx})^n - \theta(ghS_{fx})^{n+1}, \tag{63}$$

$$S_{fy} = -(1-\theta)(ghS_{fy})^n - \theta(ghS_{fy})^{n+1}, \tag{64}$$

where $n$ indicates the time level in which we know the values of the variables and $n+1$ is the next time level where we update the variables. $\theta$ is a parameter that accounts for the implicitness of the treatment of the source terms in the equation and can take any value in the interval [0,1].

So, the final expression for the numerical scheme is

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{A_i}\left(\sum_{k=1}^{NE}\left(\mathbf{E}_k^* \cdot \mathbf{n}_k ds_k - A_i \mathbf{R}_k^{1*}\right)\right)_i^n + \Delta t\left(\mathbf{R}^2\right)_i^n . \tag{65}$$

The stability criterion adopted has followed the usual in explicit finite volumes for the homogeneous system of equations not including source terms. In practice, some restrictions on the CFL can be observed due to the non-linearity of the system of equations or to the presence of source terms. Theoretical studies on this question are still on development.

$$\delta t \leq min\left[\frac{d_{ij}}{2\left(\sqrt{u^2+v^2}+c\right)_{ij}}\right], \tag{66}$$

where $d_{ij}$ is the distance between the centroid of the cell $i$ and its neighbours $j$.

## 6 2D boundaries

### *6.1 Wetting and drying fronts*

The wetting front advance over a dry bed is a moving boundary problem in the context of a depth averaged two dimensional model. As such, the optimum way to deal with it is to find the physical law that best defines the dynamics of the advancing front to use it as physical boundary condition to be plugged into the above procedure. The question about that physical law makes us reconsider the 3D basic equations at the wetting front position. In advance over adverse dry bed the water column tends to zero smoothly and, hence, the free surface and bottom level tend to reduce to one point where both the free surface and bottom boundary conditions apply simultaneously. This line of reasoning, being interesting, does not solve the discrete problem in a simple way but, on the contrary, leads to the generation of an alternative technique for a number of cells that increases in time as the wetting progresses.

In a different approach closer to the discrete solution, wetting fronts over dry surfaces can be reduced to Riemann problems in which one of the initial depths is zero. This problem can be analytically studied for simplified conditions and the solution exists both for horizontal bed (Ritter solution) [14] and for sloping bed [16]. The solution in the latter case, when dealing with adverse slopes, identifies a subset of conditions incompatible with fluid motion (stopping flow). On the other hand, the numerical technique described in section 3.1 is an approximate Riemann solver adapted to cope with zero depth cells which provides a discrete solution to the problem in all these cases not identifying correctly the stopping flow conditions. Therefore, this technique is unable to solve correctly situations of still water in a domain of irregular shape, generating spurious velocities in the wet/dry contour and often violating mass conservation. The steady flow problem is converted into an unsteady one by predicting the appearance of negative depths at the outside of the wetted domain and producing movement in water that should be always at steady state and mass conservation is lost.

In order to avoid the numerical error, the technique proposed is to enforce the local redefinition of the bottom level difference at the interface to fulfill a zero velocity equilibrium condition and therefore mass conservation. We shall call this the *wetting/drying condition* [26]. In unsteady cases, i.e., for wetting fronts advancing over an adverse dry slope, the procedure followed is the same. However in this case the numerical representation of the slope between the two adjacent cells may produce a too fast propagation of the front. It is necessary to reduce to zero the velocity components $u, v$ at the wetting/drying interface; otherwise some water could easily jump to the dry upper cell.

Previous works on this topic have reached to this point and some authors working with finite elements solve the problem allowing the controlled use of negative depths [27],

[28].

### *6.2 Boundary conditions*

The boundaries of the two-dimensional domain in which a numerical solution of the overland flow problem is sought are the different parts of the external contour line of the field. As in any other boundary problem in computational fluid dynamics, there is first a question concerning the number of physical boundary conditions required at every boundary point. To help, the theory of characteristics in 2D tells us that, depending on both the value of the normal velocity through the boundary

$$\mathbf{u} \cdot \mathbf{n} = un_x + vn_y \tag{67}$$

and the local Froude number $Fr = \mathbf{u} \cdot \mathbf{n}/c$, the possibilities are

- Supercritical inflow: $\mathbf{u} \cdot \mathbf{n} \leq -c, \quad \Rightarrow$ all the variables must be imposed.
- Subcritical inflow: $-c < \mathbf{u} \cdot \mathbf{n} \leq 0, \quad \Rightarrow$ two variables must be imposed.
- Supercritical outflow: $\mathbf{u} \cdot \mathbf{n} > c, \quad \Rightarrow$ none of the variables must be imposed.
- Subcritical outflow: $0 < \mathbf{u} \cdot \mathbf{n} \leq c, \quad \Rightarrow$ one variable must be imposed.

A second question is related to the procedure used to obtain numerical boundary conditions [3]. The idea of using a Riemann solver to calculate the flux at the face of a cell can also been used at the boundaries. The variables are stored at the centre of each cell and the boundary conditions are also imposed there. The value of the variables not prescribed are calculated from a usual finite volume balance. For this purpose, the fluxes across the edges lying on the boundary are estimated by means of a 'ghost' cell outside. Usually, the ghost cell just duplicates the boundary cell. When the boundary is a solid wall, the ghost cell is a mirror cell in which the depth of water has the same value that the boundary cell and the velocities are the same with opposite sign.

## 7 1D applications

### *7.1 Dam break test case*

The idealized dambreak problem was chosen because it is a classical example of non-linear flow with shocks to test conservation in numerical schemes and, at the same time, has an analytical solution [14]. This problem is generated by the homogeneous one-dimensional shallow water equations for the ideal case of a flat and frictionless channel of unit width and rectangular cross section, with the initial conditions

$$Q(x,0) = 0, \qquad h(x,0) = \begin{cases} h_L & \text{if } x \leq \frac{L}{2} \\ h_R & \text{if } x > \frac{L}{2} \end{cases}$$

If the calculation times used are so as to avoid interaction with the boundaries of the channel, the boundary conditions are trivial.

This classical test case is considered a benchmark for comparison of the performance of numerical schemes specially designed for discontinuous transient flow. Although defined by the system of homogeneous shallow water equations, it is widely considered a standard test case for validation of schemes. Starting from initial conditions given by still water and two different water levels separated by a dam, the theory of characteristics supplies an exact evolution solution [14] that can be used as a reference. In the example presented, two ratios of initial water depths $h_L/h_R = 10$ and $h_L/h_R = 100$ are used. The solution is displayed in Figure 2 for $t = 20$ s. An space interval of $\Delta x = 1$ m is used in the mesh. The entropy correction produces remarkable results, being the typical *"dog-leg"* effect negligible. It is also remarkable that the Lax-Wendroff scheme only with entropy correction, although displaying numerical oscillations, is able to solve strong shocks without a TVD correction. The first order upwind scheme provides a reasonably good result with a slight numerical diffusion. The second order in space TVD scheme tends to produce antidiffusive solutions, being this excessive with the Superbee flux limiter. Nevertheless with the Minmod flux limiter this is less noticeable providing a slight improvement with regard to the first order scheme. Second order in space and time improves the numerical solution being the most accurate scheme as expected.

### *7.2 Application to river flow*

In order to show the application to a practical case, an example of unsteady flow in a river is presented now. It is a 9000 m long reach of the upstream part of river Neila in Spain. Being a mountain river, it is characterized by strong irregularities in the cross section, by a rather steep part in the first kilometers and by a low base discharge (1 $m^3/s$) which, altogether, produce a high velocity basic flow, transcritical in some parts. The bottom level and the breadth variations along the axis of this river (Figure 3), together with the small value of its base discharge ($Q = 1$ $m^3/s$) and the corresponding initial values of the Froude number (Figure 4) rendered the numerical computation really challenging and accentuated the differences among the capabilities of the numerical schemes. The simulation starts from a definition of the topographic features in the form of numerical functions (data tables) of the depth and the distance along the river. These are matrices of dimensions (number of cross sections x number of water depth levels). The data tables do not correspond in general to equidistributed points along the river and never to the computational grid positions. One option is the numerical generation of intermediate sections by interpolation of the surveyed cross sections. On the other hand, the particular value of any of the functions at a nodal position for an arbitrary water

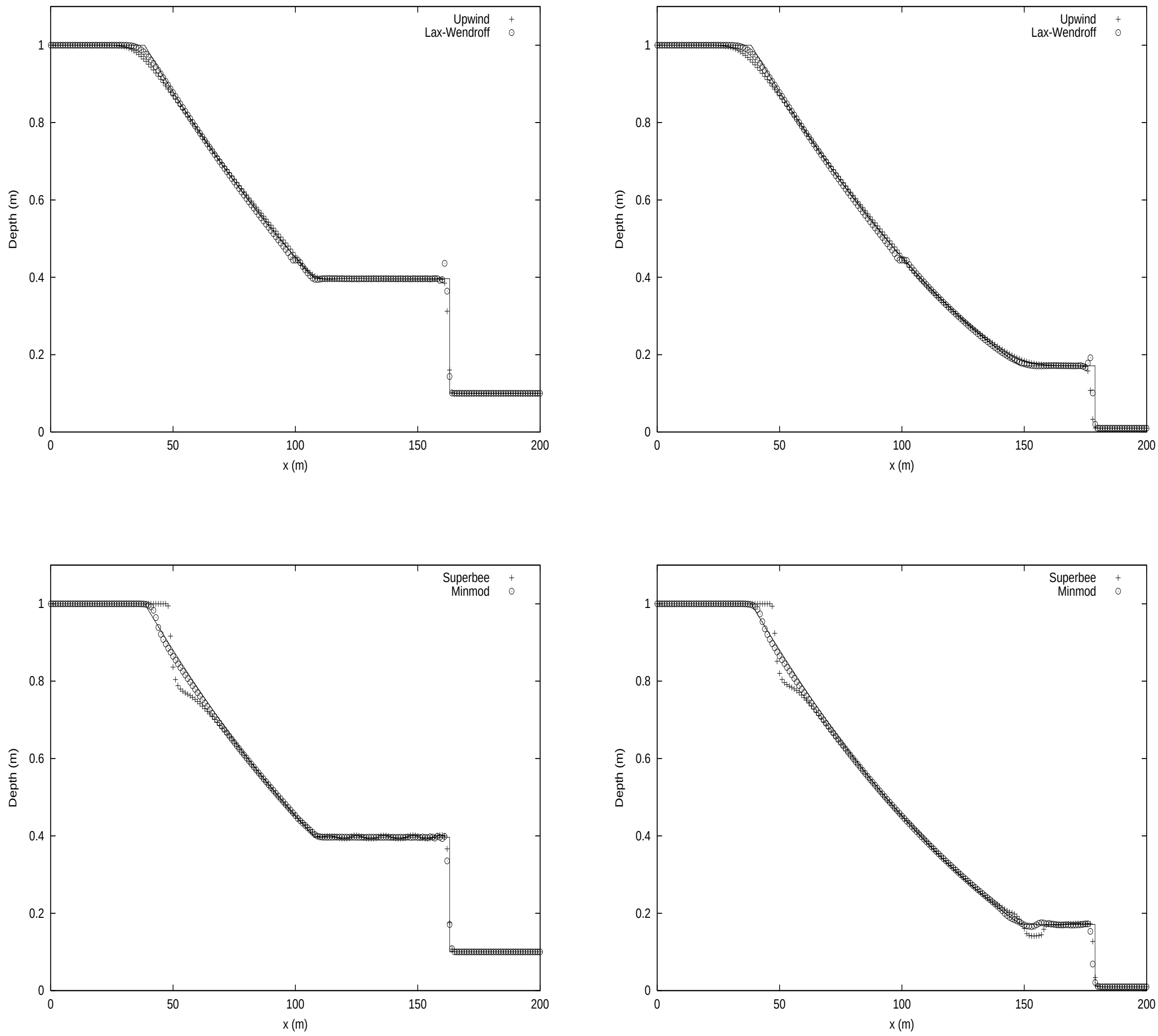


Figure 2.— Ideal dam break problem. T=20s.

depth can be obtained each time via interpolation from the data tables. More or less sophisticated interpolations can be performed. Considering that any interpolation may introduce numerical errors in the results and the irregular character of the river, a linear interpolation between data tables was used in this work. The bed slopes were determined using the values of the bottom level at each cross section.

In order to check the conservation properties of the schemes applied, and the absence of oscillations in the TVD schemes, a sudden increase in discharge to 40 $m^3/s$ and a critical depth is imposed at the upstream end. This step hydrograph propagates into the river. The same $CFL$ number as the steady flow cases and an interval of $\Delta x = 22.5$ m in the mesh are used. Figure 5 shows that the discharge wave propagates with almost a perfectly constant value at times $t = 500$ s, $t = 1000$ s and $t = 1500$ s. Fig. 6.18 shows the detail of the front wave where the advantages of using higher order approaches are noticeable, this is not so clear when reproducing steady states. In Fig. 6.19 the strong gradient in the bed slope of river Neila can be seen. Figure 6 shows some other variables as calculated with the second order in space and time TVD scheme with Superbee limiter (the most accurate scheme) and the strong irregularities of the river are evident.

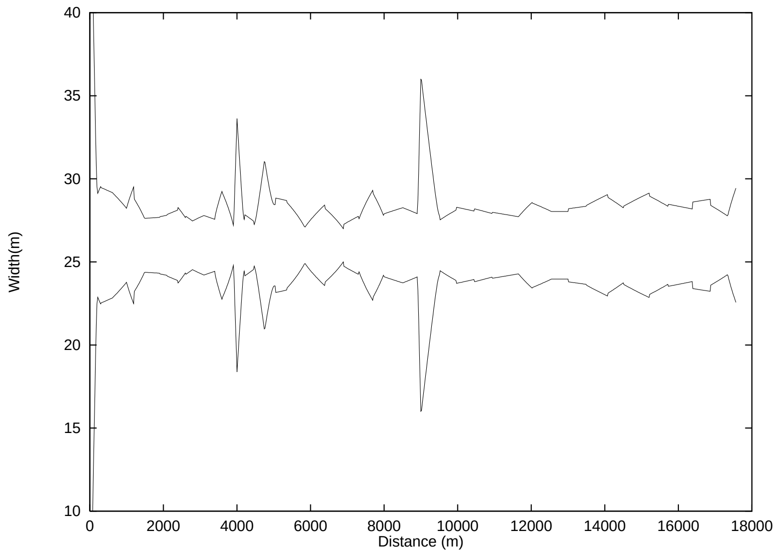


Figure 3.— Initial top width. Neila river.

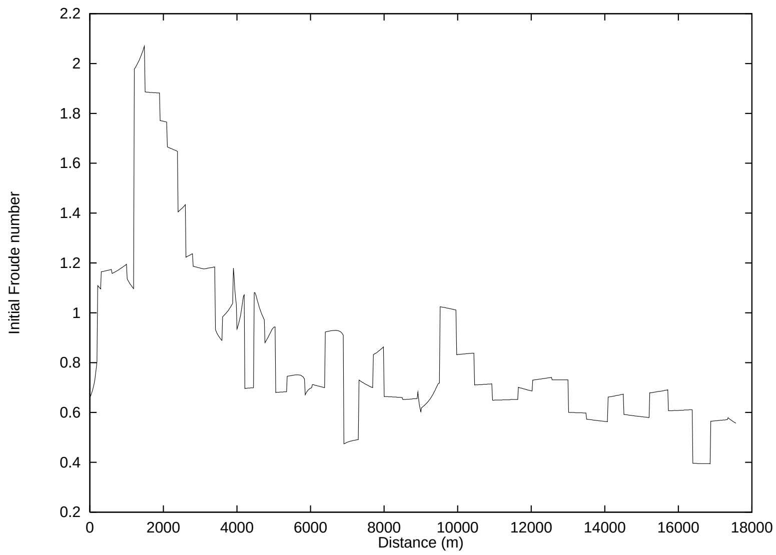


Figure 4.— Initial Froude number distribution. Neila river.

## 8 2D applications

### *8.1 Circular dam break over a dry irregular bed*

In order to test the sensitivity of the results to the treatment of the wetting/drying fronts, an academic test of a circular dam break over a highly irregular bed domain (Figure 7) is performed. The dimensions are 100x100 m and the initial water level is 30 m high, centered in the mesh, with a diameter of 20 m. The roughness coeficient of Manning, n is set equal to 0.9 in order to slow the process as much as possible.

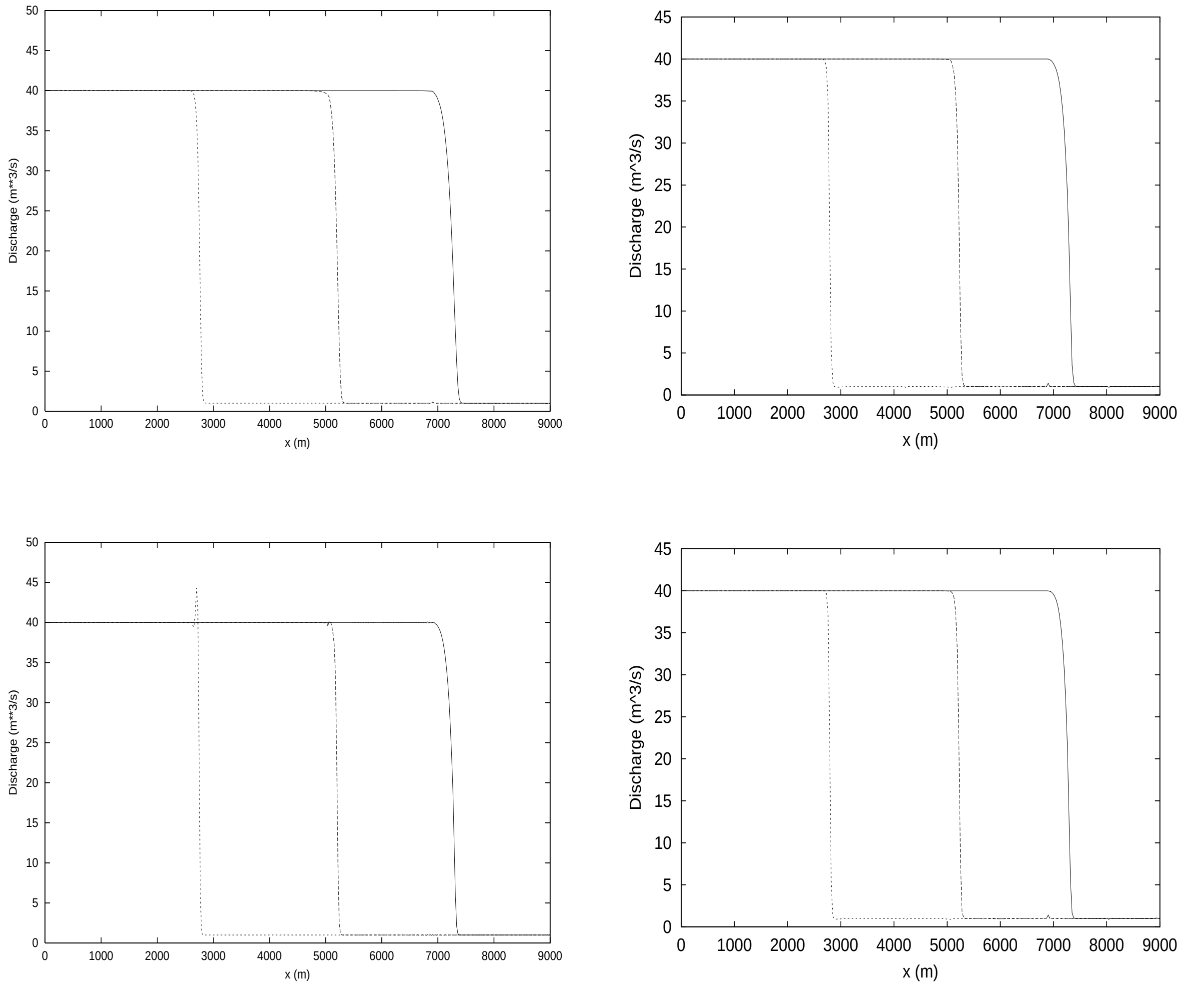


Figure 5.— Evolution in the discharge with first order upwind scheme (top left), Lax-Wendroff scheme (top right), second order in space TVD scheme with Minmod limiter (bottom left), second order in space and time TVD scheme with Superbee limiter (bottom right).

The dam break flow simulation is performed in two ways: 1) using the scheme proposed in [26] that controls the wetting/drying fronts, and 2) making zero the negative values of depth obtained when nothing is done to control the wetting/drying front. Figure 8 shows the results at different times $t = 2, 6, 12$ s, and when the steady state is achieved, ($t = 1000$ s in case 1). When option 1) is used, the still water steady state is achieved with a mass error to machine accuracy. In the case of option 2), during the phase of advance of the front, a considerable mass error is produced, and finally all the mass disappears. Figure 9 shows the results in both cases at time 2 s.

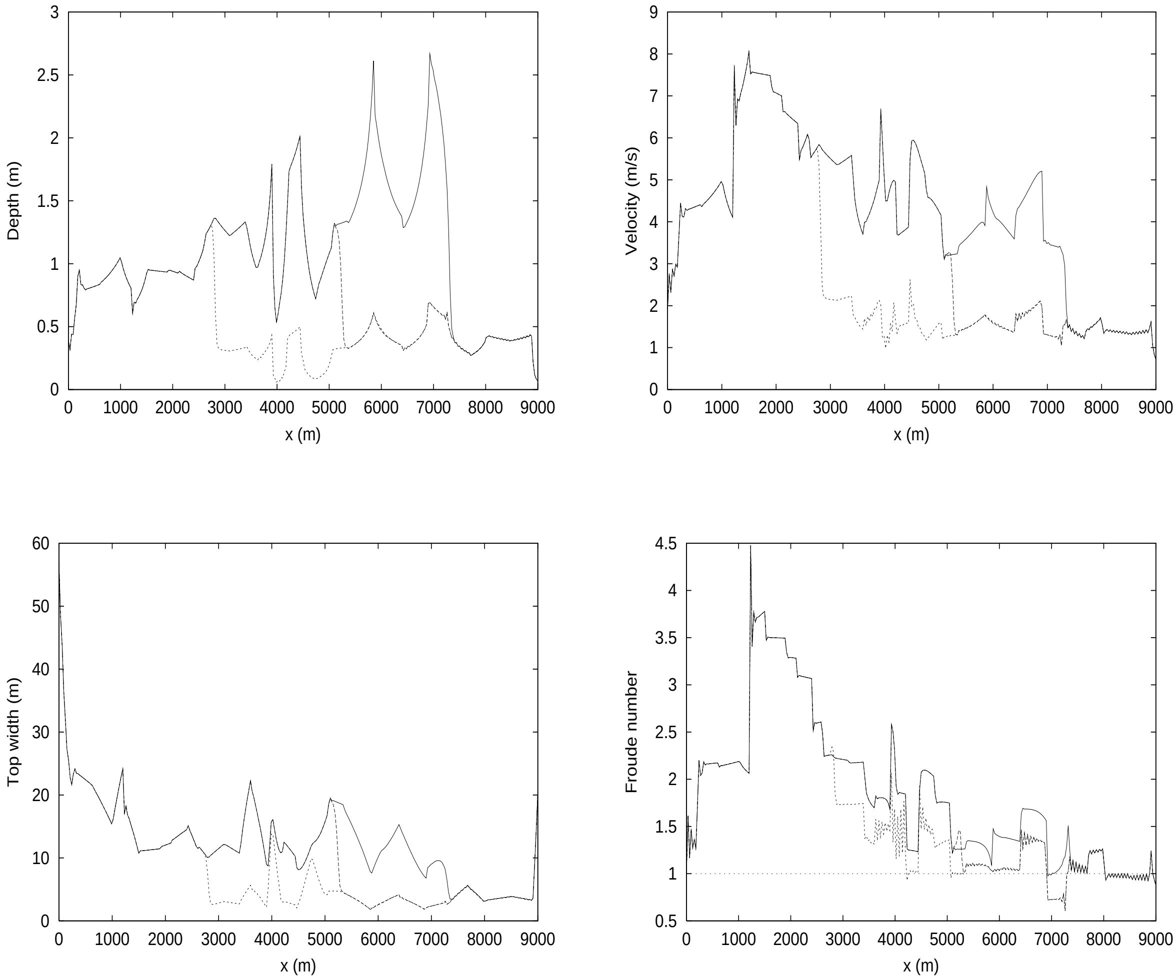


Figure 6.— Evolution in the depth, top width, velocity and Froude number with second order in space and time TVD scheme with Superbee limiter.

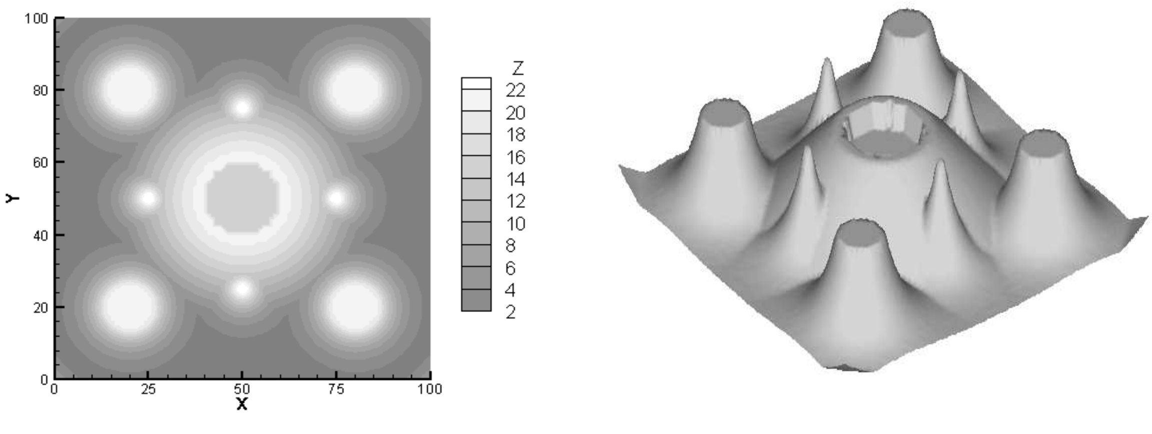


Figure 7.— Irregular bed for the circular dam break test case. Left: Plan view. Right: 3D view.

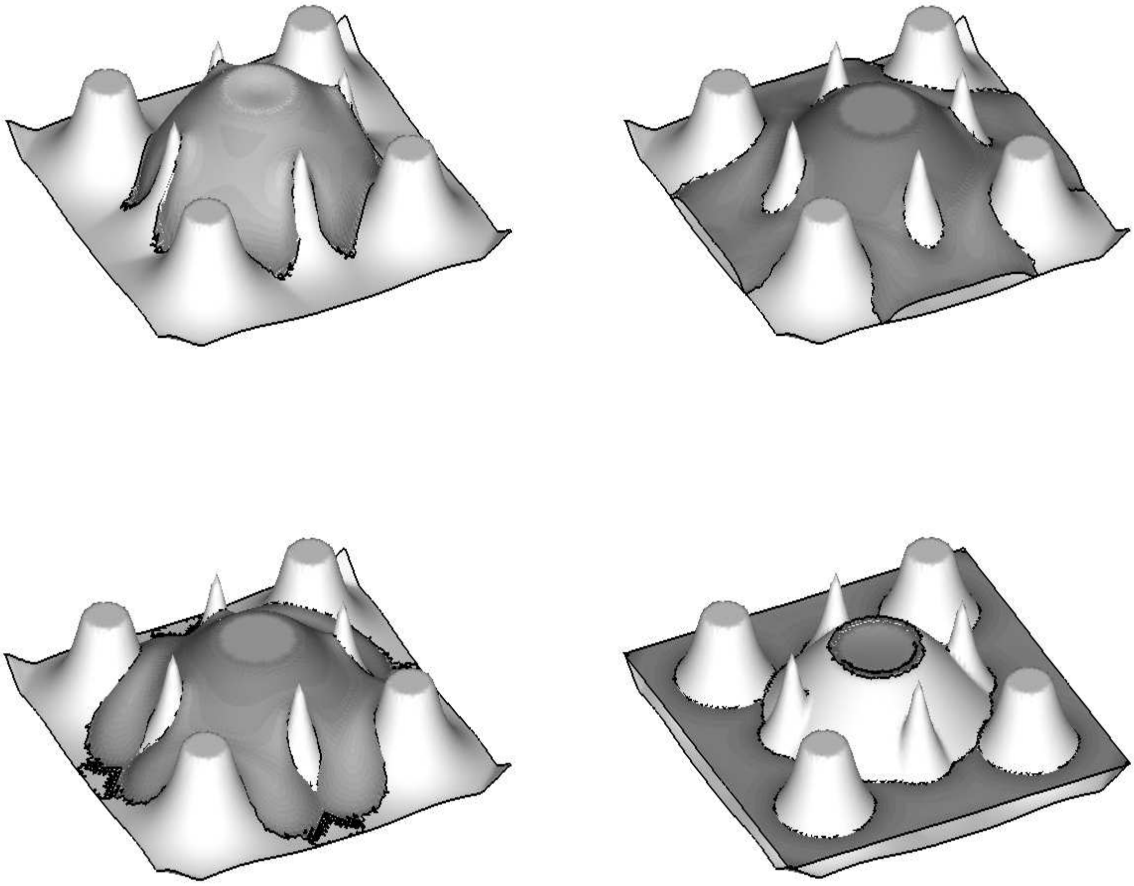

Figure 8.— Four snapshots of the water level evolution calculated with option 1). $t = 2$ s top left, $t = 6$ s top right, $t = 12$ s bottom left, and still water bottom right.

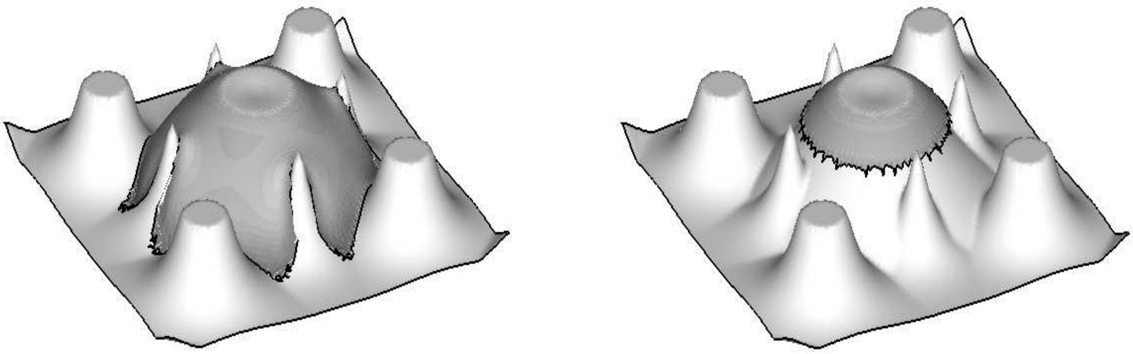

Figure 9.— Comparison of the numerical solutions provided by options 1) and 2) at $t = 2$ s after the dam breaking. Left: conservative solution. Right: non-conservative solution.

## References


[1] Abbott M.B., *Computational Hydraulics*, Ashgate, Worcester, 1992.

[2] Cunge, J.A., Holly, F.M. and Verwey, A., *Practical Aspects of Computational River Hydraulics*, Pitman, London, U.K., 1980.

[3] Hirsch,C., *Numerical Computation of Internal and External Flows*, Vol. 2, John Wiley and Sons,1990.

[4] Roe, P.L., "Approximate Riemann Solvers, Parameter Vectors and Difference Schemes" *Journal of Computational Physics* **43**, pp.357-372, 1981.

[5] Glaister, P., "Approximate Riemann solutions of the shallow water equations" *Journal of Hydraulic Research* **26** (No.3), pp. 293-306, 1988.

[6] Alcrudo, F. and García-Navarro, P., "Flux difference splitting for 1D open channel flow equations" *Int. J. for Numerical Methods in Fluids* **14**, 1992.

[7] Sweby, P.K., "High resolution schemes using flux limiters for hyperbolic conservation laws" *SIAM Journal of Numerical Analysis* **21**, 1984.

[8] LeVeque, R.J., "Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm", *J. Comput. Phys.* **146**(1):346-365, 1998.

[9] Bermúdez, A. and Vázquez, M.E., "Upwind methods for hyperbolic conservation laws with source terms" *Computers and Fluids* **8** 1049-1071, 1994.

[10] Barley, J.J., "A Survey of Operator Splitting Applied to Upwind Differencing", *Numerical Analysis Report* **5**, University of Reading, 1988.

[11] Katopodes, N., "Two-dimensional Surges and Shocks in Open Channels." *J. of Hydraulic Engineering, ASCE* **110**, pp. 794-812, 1984.

[12] Katopodes, N., and Strelkoff, T., "Computing Two-dimensional Dam-break Flood Waves." *J. of Hydraulic Engineering, ASC* **104**, pp. 1269-1288, 1978.

[13] Chanson, H., *The hydraulics of open channel flow,* Arnold, 1999.

[14] Stoker, J.J., *Water Waves*, Interscience, New York, 1957.

[15] Vreugdenhil, C.B., *Numerical methods for shallow-water flow.* Kluwer Ac. Pub., Dordrecht, The Netherlands, 1994.

[16] Toro, E.F., *Shock Capturing Methods for Free-Surface Shallow Flows*, John Wiley and Sons (2001).

[17] Burguete, J., García-Navarro, P.: "Improving simple explicit methods for unsteady open channel and river flow", *International Journal for Numerical Methods in Fluids* **45**, pp. 125-156, 2004.

[18] Burguete, J., García-Navarro, P. "Efficient construction of high-resolution TVD conservative schemes for equations with source terms: application to shallow water flows", *International Journal for Numerical Methods in Fluids* **37**, pp.209-248, 2001.

[19] García-Navarro, P. and Vázquez-Cendón, M.E., "On numerical treatment of the source terms in the shallow water equations", *Computers and Fluids* **29**, pp. 951-979, 2000.

[20] Vázquez-Cendón, M.E., "Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry", *Journal of Computational Physics* **148**, 1999.

[21] Lax, P.D., Wendroff, B. "Systems of conservation laws", *Comm. Pure and Applied Mathematics* **13** pp. 217-237, 1960.

[22] LeVeque, R. J., *Numerical methods for conservation laws*, Birkhauser Basel, 1992.

[23] Roe, P.L., "A Basis for Upwind Differencing of the Two-Dimensional Unsteady Euler Equations", in *Num. Met. Fluid Dyn. II*, Oxford Univ. Press, 1986.

[24] Brufau, P. and García-Navarro, P., "Two-dimensional dam break flow simulation", *Int. J. for Num. Meth. in Fluids* **33**, pp. 35-57, 2000.

[25] Bermúdez, A., Dervieux, P., Desideri, J.A. and Vázquez, M.E., "Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes", *Comput. Methods Appl. Mech. Engrg.* **155**, pp. 49-72, 1998.

[26] Brufau, P., Vazquez-Cendón, M.E. and García-Navarro. P., "A numerical model for the flooding and drying of irregular domains", *Int. J. for Num. Meth. in Fluids* **39**, pp. 247-275, 2002.

[27] Sleigh, P.A., Berzins, M., Gaskell, P.H. and Wright, N.G., "An Unstructured Finite Volume Algorithm for predicting flow in rivers and stuaries", in *Comp. and Fluids*. 1997.

[28] Khan, A.A., "Modelling flow over an initially dry bed", *J. of Hyd. Res.* **38**(5), pp. 383-389, 2000.