

Extension of an explicit finite volume method to large time steps (CFL > 1): application to shallow water flows

J. Murillo[‡], P. García-Navarro^{*†}, P. Brufau[§] and J. Burguete[¶]

Fluid Mechanics, CPS, University of Zaragoza, Zaragoza, Spain

SUMMARY

In this work, the explicit first order upwind scheme is presented under a formalism that enables the extension of the methodology to large time steps. The number of cells in the stencil of the numerical scheme is related to the allowable size of the CFL number for numerical stability. It is shown how to increase both at the same time. The basic idea is proposed for a 1D scalar equation and extended to 1D and 2D non-linear systems with source terms. The importance of the kind of grid used is highlighted and the method is outlined for irregular grids. The good quality of the results is illustrated by means of several examples including shallow water flow test cases. The bed slope source terms are involved in the method through an upwind discretization. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: large CFL; explicit schemes; unstructured grids; shallow water flow; finite volumes; upwind method

0. INTRODUCTION

Upwind methods have proved very successful in computational fluid dynamics (CFD), mainly in connection with aerodynamics where they have gained widespread acceptance [1]. They are, perhaps, the most widely researched algorithms in connection with unsteady flow simulation codes. At the same time, numerical methods to predict water profile, velocities and discharges in hydraulic systems modelling have become a common tool. Upwind methods in particular are becoming increasingly popular in the hydraulics literature and have proved a suitable way to discretize the shallow water equations [2–5]. Being a hyperbolic system of conservation laws, they are a good candidate for application of the techniques developed for the Euler equations in gas dynamics.

*Correspondence to: Pilar García-Navarro, Fluid Mechanics, CPS, University of Zaragoza, Zaragoza, Spain.

†E-mail: pigar@unizar.es

‡E-mail: jmurillo@mafalda.cps.unizar.es

§E-mail: brufau@unizar.es

¶E-mail: jburguete@able.es

Received 15 November 2004

Revised 17 May 2005

Accepted 24 May 2005

Finite difference schemes for time-dependent advective equations are traditionally divided in two main groups, according to the way of discretization used for the time derivative, as explicit and implicit. Implicit schemes offer numerical stability (not always unconditional, however) at the extra cost of having to deal with the resolution of an algebraic, and often non-linear, system with as many unknowns as grid points at every time step. On the other hand, conceptual simplicity is the most valuable characteristic of the former in which variables at a future time can be independently evaluated at every single point. The allowable time step size is nevertheless restricted in the explicit case by stability reasons to fulfil the Courant–Friedrichs–Lewy (CFL) condition [6].

It is possible to relax the condition over the time step size when using explicit schemes. A generalization of the first order explicit upwind and Roe’s method, modified to allow large time steps, was explored by Leveque [7, 8] first in the scalar non-linear case and then adapted to systems of equations. It becomes stable for CFL’s larger than 1 and provides an accurate and correct solution of shocks. In the case of a linear problem, it is reduced to a linear interpolation scheme. The technique is devised to cope well with flow transients and even with discontinuities far from the boundaries, being able to give a resolution of the shocks even sharper for $CFL > 1$. Although the technique approximates linearly non-linear interactions, it is able to correctly reproduce the discontinuities of interest in common hydraulic calculations allowing the use of time steps not restricted by the usual CFL stability conditions. Other approaches have been envisaged and applied to different non-linear problems in References [9–12].

First in this paper, we shall describe the algorithm for the 1D scalar case as well as for 1D systems of equations. The performance of the scheme is evaluated to solve the shallow water equations first for the ideal dam break problem because it is a classical example of non-linear flow where shocks and rarefactions appear and, at the same time, has an exact solution. The way to deal with bed slope and friction source terms and to incorporate them into the proposed procedure will be also presented starting by the 1D approach. A problem with analytical solution of steady open channel flow with bed slope and friction is used in this part as validation test case.

In the second part of the work the method will be extended to a 2D finite volume approach. The additional difficulties met when adapting it to unstructured meshes will be identified pointing out the importance of the grid irregularity. Apart from a 2D scalar advection example, the extended time step explicit scheme will be applied to shallow water problems of different types including laboratory test cases with experimental data as well as a steady problem over variable bottom with analytical solution and transient flow over irregular bed.

1. NUMERICAL SCHEME. ONE-DIMENSIONAL CASE

1.1. Scalar equation

The numerical resolution of the scalar differential equation

$$u_t + f(u, x)_x = 0, \quad \lambda = f_u \quad (1)$$

in a fixed regular discrete mesh $\{x_i, i = 1, N\}$, letting aside grid-adaption considerations [13], by means of a first order upwind scheme with forward Euler time stepping, is based on a

piecewise constant approximation of the function u . The value of the average discrete celerity $\tilde{\lambda}_{i+1/2}$ at each interface $(i, i + 1)$ defines the flux differences as contributions that may emerge from the interface according to the sign of $\tilde{\lambda}_{i+1/2}$:

$$\delta f_{i+1/2}^{\pm} = \tilde{\lambda}_{i+1/2}^{\pm} \delta u_{i+1/2} \quad (2)$$

so that the function at each cell node i is updated in one time step Δt , according to the incoming contributions from the left or right interfaces

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (\delta f_{i-1/2}^+ + \delta f_{i+1/2}^-) \quad (3)$$

where Δx is the regular length of the spatial interval between nodes. Usually, a numerical flux is defined and the above is re-expressed as

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (f_{i+1/2}^* - f_{i-1/2}^*) \quad (4)$$

with

$$f_{i+1/2}^* = \frac{1}{2}(f_{i+1} + f_i) - \frac{1}{2} |\tilde{\lambda}_{i+1/2}| \delta u_{i+1/2} \quad (5)$$

This is a finite volume point of view that updates the value of the function at the cell as a result of the net flux through the cell edges. There is an alternative way of considering the situation as centred at the interfaces and looking where the contributions go. Always at the interface $(i, i + 1)$, a signal $v_{i+1/2} \delta u_{i+1/2} = \frac{\Delta t}{\Delta x} \tilde{\lambda}_{i+1/2} \delta u_{i+1/2}$ is defined and the rule to follow is [7]

$$\begin{aligned} \text{if } \tilde{\lambda}_{i+1/2} > 0 &\Rightarrow v_{i+1/2} \delta u_{i+1/2} \text{ updates } i + 1 \\ \text{if } \tilde{\lambda}_{i+1/2} < 0 &\Rightarrow |v_{i+1/2}| \delta u_{i+1/2} \text{ updates } i \end{aligned} \quad (6)$$

where $v_{i+1/2} = (\Delta t / \Delta x) \tilde{\lambda}_{i+1/2}$ is a dimensionless coefficient.

Both approaches (3) and (4) are equivalent and equally simple for the case $\text{CFL} < 1$, where in this case

$$\text{CFL} = \frac{\Delta t}{\Delta t_{\max}}, \quad \Delta t_{\max} = \min \left\{ \frac{\Delta x}{|\tilde{\lambda}_{i+1/2}|} \right\}_{i=1, \text{NCELL}} \quad (7)$$

being NCELL the total number of cells. This stability condition is related to the maximum amount of information that the pair of cells in the stencil of the scheme can exchange in one time step. The second approach as in (6) is preferable to extend the scheme to $\text{CFL} > 1$. As described by Leveque [7, 8], in the case of using a large time step, more cells get involved. The interface contribution that defines the basic explicit conservative method (3) is split in several pieces, here called waves, depending on the size of the time step and these waves are transported to the cells according to the following algorithm:

When $\tilde{\lambda}_{i+1/2} > 0$

$$\begin{aligned} \delta u_{i+1/2} &\text{ updates } i + 1, \dots, i + \mu_{i+1/2} \\ (v - \mu)_{i+1/2} \delta u_{i+1/2} &\text{ updates } i + \mu_{i+1/2} + 1 \end{aligned} \quad (8)$$

and when $\tilde{\lambda}_{i+1/2} < 0$

$$\begin{aligned} \delta u_{i+1/2} & \text{ updates } i, \dots, i + \mu_{i+1/2} + 1 \\ |v - \mu|_{i+1/2} \delta u_{i+1/2} & \text{ updates } i + \mu_{i+1/2} \end{aligned} \tag{9}$$

where $\mu_{i+1/2} = \text{int}(v_{i+1/2})$. When $\text{CFL} \leq 1$, the basic scheme is recovered and it is important to note that in (8) and (9) the dimensionless coefficients are all smaller or equal to 1 hence ensuring that the solution will not be amplified. The scheme so constructed is explicit and conservative in the same sense as (3) is.

Figure 1 shows how the information is distributed to several cells when $\tilde{\lambda}_{i+1/2}$ is negative (left) and when $\tilde{\lambda}_{i+1/2}$ is positive (right) and both values are used to update the information in the cells. All the cell values are updated at the end of the process of distribution of the different contributions.

1.1.1. Boundary treatment. In open boundaries, the boundary cell receives the information furnished by the neighbour cells according to the scheme provided in (8) or (9). Some of the contributions cross the boundaries and do not affect the updated solution of the boundary cell. The next example is intended to clarify how, despite the method does not use the signals that cross the downstream boundary, the method remains conservative and the final solution is correct. Figure 2 displays a simple uniform grid of only three cells, one interior and two boundary cells. The initial function u is advected with a constant and uniform velocity $\lambda > 0$, from time t^n to time $t^{n+2} = t^n + 2\Delta t$. The upstream condition is $u(1, t) = u_a$. On the upper part of the figure, the left column (a), shows the procedure in the two time steps using $\text{CFL} = 1$ whereas the right column (b), shows the procedure in one single time step using $\text{CFL} = 2$. In the lower part of the figure, (c) displays the solution reached by the two calculations at the final time. It is clear that, when $\text{CFL} = 1$, two steps are involved and both are conservative:

$$(\delta u_{i+1/2}^n + \delta u_{i+3/2}^n) \Delta x = \lambda(u_1^n - u_3^n) \Delta t \Rightarrow \delta u_{i+1/2}^n + \delta u_{i+3/2}^n = u_1^n - u_3^n \tag{10}$$

$$(\delta u_{i+1/2}^{n+1} + \delta u_{i+3/2}^{n+1}) \Delta x = \lambda(u_1^{n+1} - u_3^{n+1}) \Delta t \Rightarrow \delta u_{i+1/2}^{n+1} + \delta u_{i+3/2}^{n+1} = u_1^{n+1} - u_3^{n+1} \tag{11}$$

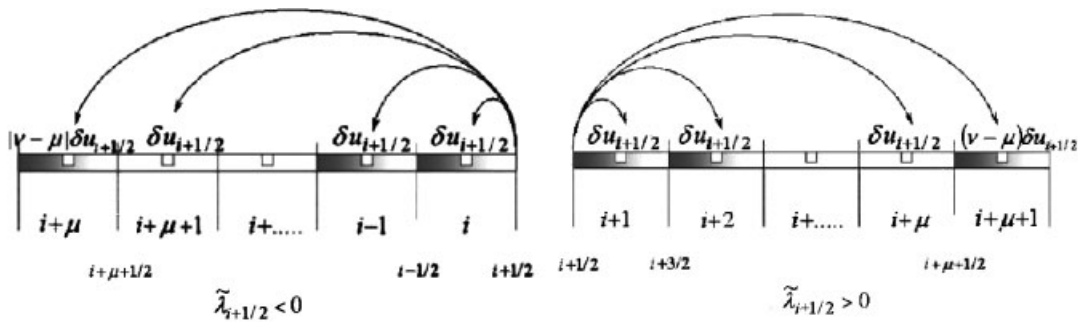


Figure 1. Distribution of the signals for different signs of the discrete velocity.

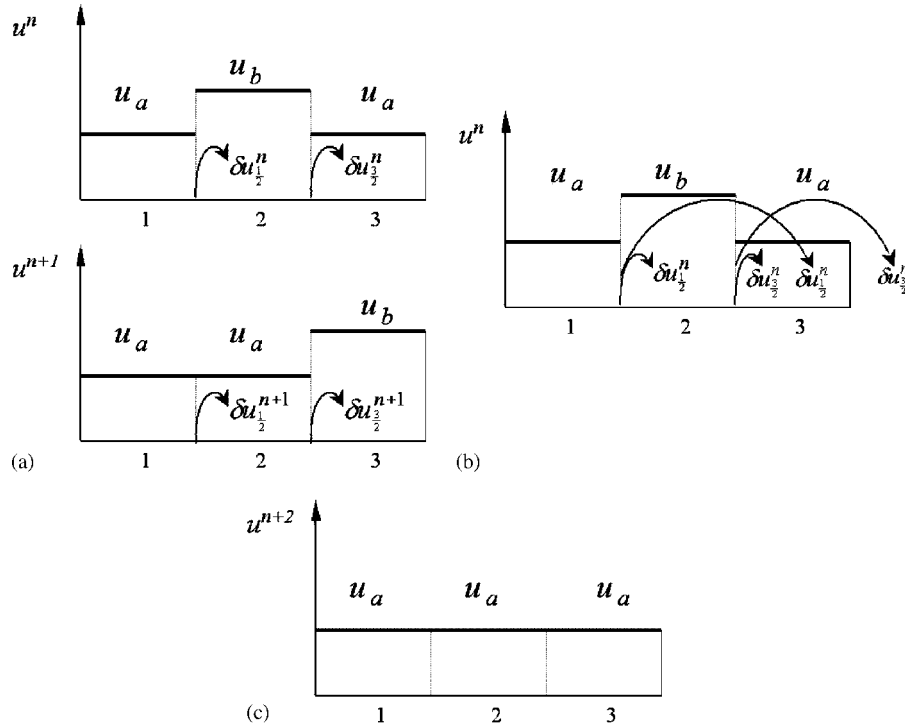


Figure 2. (a) Signal distribution with CFL=1; (b) signal distribution with CFL=2; and (c) solution at $t^{n+2} = t^n + 2\Delta t$.

When using CFL=2, only one conservative step is used:

$$(2\delta u_{i+1/2}^n + 2\delta u_{i+3/2}^n)\Delta x = \lambda(u_1^n - u_3^n)2\Delta t \Rightarrow \delta u_{i+1/2}^n + \delta u_{i+3/2}^n = u_1^n - u_3^n \tag{12}$$

In the case of zero flow or closed boundaries, an accumulation technique is proposed. Let us consider first the upstream boundary and the contributions travelling from interface $i + 1/2$ ($\tilde{\lambda}_{i+1/2} < 0$). If the value of $i + \mu$ is less than 1, it means that some of the signals from cell edge $i + 1/2$ would go out of the upstream end of the domain. As the solid wall condition requires that no information crosses the boundary and the method must remain conservative, all these contributions are accumulated at the upstream boundary cell. Therefore, in this case, the value of $\delta u_{i+1/2}$ is stored at cell 1, $|v - i + 1|$ times. On the other hand, the contributions arriving from interface $i + 1/2$ to the downstream cell N are considered ($\tilde{\lambda}_{i+1/2} > 0$). When the value of $i + \mu + 1$ exceeds the number of the last cell, N , the quantity $\delta u_{i+1/2}$ must be stored at cell N , $i + 1 + v - N$ times. Otherwise, the method does not remain conservative. Both situations are sketched in Figure 3.

1.1.2. Irregular mesh. When using an irregular mesh things are different. Considering Δx_i the length of the cell centred at node i , the finite volume formulation of (3) is

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x_i} (\delta f_{i-1/2}^+ + \delta f_{i+1/2}^-) \tag{13}$$

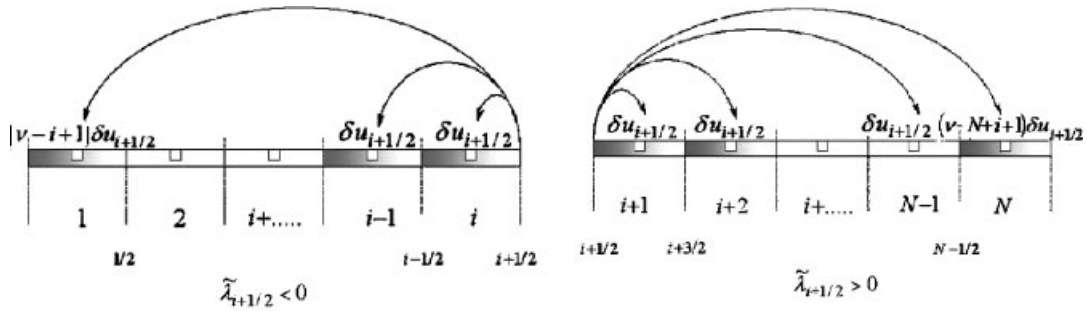


Figure 3. Accumulation of signals at the upstream (left) and downstream (right) boundaries.

and the local or cell edge dimensionless number is redefined as

$$v_{i+1/2} = \frac{\Delta t}{\Delta x_{i+s}} \tilde{\lambda}_{i+1/2} \quad s = \frac{1}{2}(1 + \text{sign}(\tilde{\lambda}_{i+1/2})) \quad (14)$$

so that the above procedure is extended following:

if $\tilde{\lambda}_{i+1/2} > 0$

$$\begin{aligned} \delta u_{i+1/2} & \text{ updates } i + 1 \\ \frac{\Delta x_{i+1}}{\Delta x_{i+j}} \delta u_{i+1/2} & \text{ updates } i + j, \quad j = 2, \dots, \mu_{i+1/2} \\ (v - \mu)_{i+1/2} \frac{\Delta x_{i+1}}{\Delta x_{i+\mu+1}} \delta u_{i+1/2} & \text{ updates } i + \mu_{i+1/2} + 1 \end{aligned} \quad (15)$$

else if $\tilde{\lambda}_{i+1/2} < 0$

$$\begin{aligned} \delta u_{i+1/2} & \text{ updates } i \\ \frac{\Delta x_i}{\Delta x_{i+j}} \delta u_{i+1/2} & \text{ updates } i + j, \quad j = -1, \dots, \mu_{i+1/2} + 1 \\ |v - \mu|_{i+1/2} \frac{\Delta x_i}{\Delta x_{i+\mu}} \delta u_{i+1/2} & \text{ updates } i + \mu_{i+1/2} \end{aligned} \quad (16)$$

where again $\mu_{i+1/2} = \text{int}(v_{i+1/2})$. In order to ensure the good behaviour of the numerical solution in this case, the dimensionless coefficients in (15) and (16) must always be kept smaller or equal to one. This is guaranteed provided that the definition of Δt_{\max} given in (7) is carefully generalized to the irregular case. If (7) is extended to

$$\Delta t_{\max} = \min \left\{ \frac{\Delta x_i}{|\tilde{\lambda}_{i+1/2}|} \right\}_{i=1, \text{NCELL}} \quad (17)$$

The resulting value does not ensure the stability required. Instead, a two-step search of the reference time step size is proposed. First, the reference time step size is found for every cell mesh taking into account the stencil for each CFL value

$$\Delta t_{\max,i} = \begin{cases} \min \left\{ \frac{\Delta x_i, \Delta x_{i+1}, \Delta x_{i+2}, \dots, \Delta x_{i+\mu+1}}{|\tilde{\lambda}_{i+1/2}|} \right\} & \text{if } \tilde{\lambda}_{i+1/2} > 0 \\ \min \left\{ \frac{\Delta x_{i+1}, \Delta x_i, \Delta x_{i-1}, \dots, \Delta x_{i-\mu}}{|\tilde{\lambda}_{i+1/2}|} \right\} & \text{if } \tilde{\lambda}_{i+1/2} < 0 \end{cases} \quad (18)$$

Then, the minimum value among all of them is chosen.

$$\Delta t_{\max} = \min\{\Delta t_{\max,i}\}_{i=1, \text{NCELL}} \quad (19)$$

1.2. Scalar equation with source terms

If a source term $s = s(u, x)$ is added to the scalar equation (1),

$$u_t + f(u, x)_x = s(u, x) \quad (20)$$

the problem can be discretized using a regular grid and assuming that the flux derivative and the source term can be decomposed in a similar manner [4, 14] as

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (\delta f_{i-1/2}^+ + s_{i-1/2}^+ + \delta f_{i+1/2}^- + s_{i-1/2}^-) \quad (21)$$

with

$$s_{i\mp 1/2}^\pm = \frac{1}{2}(1 \pm \text{sgn}(\tilde{\lambda}_{i\mp 1/2}^\pm))s_{i\mp 1/2} \quad (22)$$

The contributions from the source terms are also evaluated at each cell interface depending on the value of the average velocity $\tilde{\lambda}_{i+1/2}$. In order to generalize (21) to CFL > 1 numbers, the scheme is written in a more compact way:

$$u_i^{n+1} = u_i^n - (v_{i-1/2}^+ \delta u_{i-1/2} + \kappa_{i-1/2}^+ \Delta t + v_{i+1/2}^- \delta u_{i+1/2} + \kappa_{i+1/2}^- \Delta t) \quad (23)$$

where

$$v_{i\mp 1/2}^\pm = \frac{\Delta t}{\Delta x} \tilde{\lambda}_{i\mp 1/2}^\pm, \quad \kappa_{i\mp 1/2}^\pm = \frac{1}{\Delta x} s_{i\mp 1/2}^\pm \quad (24)$$

that can be rewritten as

$$\begin{aligned} u_i^{n+1} = u_i^n - (v_{i-1/2}^+ \delta u_{i-1/2} + \kappa_{i-1/2}^+ (\mu_{i-1/2} \Delta t_1 + \Delta t_2)_{i-1/2} + v_{i+1/2}^- \delta u_{i+1/2} \\ + \kappa_{i+1/2}^- (\mu_{i+1/2} \Delta t_1 + \Delta t_2)_{i+1/2}) \end{aligned} \quad (25)$$

where two new time variables, devoid of physical meaning, are defined, Δt_1 and Δt_2 , at each interface:

$$\Delta t_{1,i\pm 1/2} = \frac{\Delta x}{|\tilde{\lambda}_{i\pm 1/2}|}, \quad \Delta t_{2,i\pm 1/2} = \Delta t_1 |v - \mu|_{i\pm 1/2} \quad (26)$$

Hence, the algorithm to propagate the signals to the surrounding cells is

When $\tilde{\lambda}_{i+1/2} > 0$

$$\begin{aligned} \delta u_{i+1/2} + (\kappa \Delta t_1)_{i+1/2} & \text{ updates } i + 1, \dots, i + \mu_{i+1/2} \\ (v - \mu)_{i+1/2} \delta u_{i+1/2} + (\kappa \Delta t_2)_{i+1/2} & \text{ updates } i + \mu_{i+1/2} + 1 \end{aligned} \quad (27)$$

and when $\tilde{\lambda}_{i+1/2} < 0$

$$\begin{aligned} \delta u_{i+1/2} + (\kappa \Delta t_1)_{i+1/2} & \text{ updates } i, \dots, i + \mu_{i+1/2} + 1 \\ |v - \mu|_{i+1/2} \delta u_{i+1/2} + (\kappa \Delta t_2)_{i+1/2} & \text{ updates } i + \mu_{i+1/2} \end{aligned} \quad (28)$$

For the case $CFL < 1$, $\Delta t_2 = \Delta t_{\max}$. Therefore, the initial philosophy developed by Leveque [7] and also used by [15] is followed introducing a source term in the equation. When solid wall boundaries are present, the accumulation technique described in the above section is used including the source terms as well. The algorithm in (27), (28) must be modified to include the influence of the different cell sizes, as in (15), (16) in case of using an irregular grid.

1.3. One-dimensional systems of equations with source terms

Let us consider a hyperbolic system with source terms in the form

$$\mathbf{U}_t + \mathbf{F}_x = \mathbf{S}(x, \mathbf{U}), \quad \mathbf{A} = \frac{d\mathbf{F}}{d\mathbf{U}} \quad (29)$$

Provided that Roe's linearization is used to decouple the hyperbolic system, an approximate matrix \mathbf{A}^* can be built whose eigenvalues $\tilde{\lambda}^m$ and eigenvectors \mathbf{e}^m can be used to express the flux difference as a sum of waves [1]:

$$\begin{aligned} \delta \mathbf{U}_{i+1/2} &= \mathbf{U}_{i+1} - \mathbf{U}_i = \sum_m (\alpha^m \mathbf{e}^m)_{i+1/2} = \sum_m \delta \mathbf{U}_{i+1/2}^m \\ \delta \mathbf{F}_{i+1/2} &= \mathbf{F}_{i+1} - \mathbf{F}_i = \sum_m (\tilde{\lambda}^m \alpha^m \mathbf{e}^m)_{i+1/2} = \sum_m (\tilde{\lambda}^m \delta \mathbf{U}^m)_{i+1/2} \end{aligned} \quad (30)$$

and to discretize the source terms as well [3, 4]

$$\mathbf{S}_{i+1/2} = \sum_m (\beta^m \mathbf{e}^m)_{i+1/2} = \sum_m \mathbf{S}_{i+1/2}^m \quad (31)$$

so that, in a regular grid, the cells are updated following the scheme

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \frac{\Delta t}{\Delta x} (\delta \mathbf{F}_{i-1/2}^+ + \mathbf{S}_{i-1/2}^+ + \delta \mathbf{F}_{i+1/2}^- + \mathbf{S}_{i-1/2}^-) \quad (32)$$

where

$$\mathbf{S}_{i+1/2}^{m\pm} = \frac{1}{2} (1 \pm \text{sgn}(\tilde{\lambda}_{i+1/2}^m)) \mathbf{S}_{i+1/2}^m \quad (33)$$

Therefore, the following can be written:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \left(\sum_m (v \delta \mathbf{U})_{i-1/2}^{m+} + \sum_m \kappa_{i-1/2}^{+m} \Delta t + \sum_m (v \delta \mathbf{U})_{i+1/2}^{m-} + \sum_m \kappa_{i+1/2}^{-m} \Delta t \right) \quad (34)$$

with

$$v_{i\mp 1/2}^{m\pm} = \frac{\Delta t}{\Delta x} \tilde{\lambda}_{i\mp 1/2}^{m\pm} \quad \text{and} \quad \kappa_{i\mp 1/2}^{m\pm} = \frac{1}{\Delta x} \mathbf{S}_{i\mp 1/2}^{m\pm}$$

The basic idea is again to calculate $\delta \mathbf{U}$ at every interface and propagate the different m waves according to the sign of their celerities (eigenvalues) and the values of the local parameter. For instance, for each m component of $\delta \mathbf{U}_{i+1/2}$, $\delta \mathbf{U}_{i+1/2}^m$, the following is set:

When $\tilde{\lambda}_{i+1/2}^m > 0$

$$\begin{aligned} \delta \mathbf{U}_{i+1/2}^m + (\kappa^m \Delta t_1^m)_{i+1/2} & \text{ updates } i+1, \dots, i + \mu_{i+1/2}^m \\ (v - \mu) \delta \mathbf{U}_{i+1/2}^m + (\kappa^m \Delta t_2^m)_{i+1/2} & \text{ updates } i + \mu_{i+1/2}^m + 1 \end{aligned} \quad (35)$$

When $\tilde{\lambda}_{i+1/2}^m < 0$

$$\begin{aligned} \delta \mathbf{U}_{i+1/2}^m + (\kappa^m \Delta t_1^m)_{i+1/2} & \text{ updates } i, \dots, i + \mu_{i+1/2}^m + 1 \\ |v - \mu| \delta \mathbf{U}_{i+1/2}^m + (\kappa^m \Delta t_2^m)_{i+1/2} & \text{ updates } i + \mu_{i+1/2}^m \end{aligned} \quad (36)$$

where $\mu_{i+1/2}^m = \text{int}(v_{i+1/2}^m)$, $\Delta t_1^m = \Delta x / |\tilde{\lambda}_{i+1/2}^m|$ and $\Delta t_2^m = |v - \mu|_{i+1/2} \Delta t_1^m$.

1.4. Application to 1D shallow water flow

The 1D shallow water equations are accepted to model a certain kind of free surface flow [5]. In case of a unit width rectangular cross section open channel flow, the conservative form of the system of equations can be expressed as in (26) with

$$\mathbf{U} = \begin{pmatrix} h \\ hu \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} hu \\ hu^2 + g \frac{h^2}{2} \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} 0 \\ gh(S_0 - S_f) \end{pmatrix} \quad (37)$$

where h represents the water depth, g is the acceleration of the gravity and u is the average component of the velocity along the x direction. The source terms in the momentum equation are the bed slope along the x -axis and the friction losses

$$S_0 = -\frac{\partial z}{\partial x}, \quad S_f = \frac{n^2 Q |Q|}{A^{10/3}} P^{4/3} \quad (38)$$

where S_f is expressed in terms of the Manning's roughness coefficient n , the wetted area A , and the wetted perimeter P . The Jacobian form and properties can be found in several references [3, 4] and will not be repeated here.

The source term vector \mathbf{S} is treated in two parts: the bottom variations \mathbf{B} and the friction term \mathbf{R} , $\mathbf{S} = \mathbf{B} + \mathbf{R}$. As the slope part contains a spatial derivate an upwind approach is adopted to model the bottom variations in order to ensure the best balance with the flux terms at least in steady-state cases [4]. The friction term \mathbf{R} is discretized in a pointwise manner, so that the

final proposed expression for the numerical scheme is

$$\begin{aligned} \mathbf{U}_i^{n+1} = & \mathbf{U}_i^n - \left(\sum_{m=1}^2 (v\delta\mathbf{U})_{i-1/2}^{m+} + \sum_{m=1}^2 \mathbf{k}_{i-1/2}^{m+} \Delta t \right. \\ & \left. + \sum_{m=1}^2 (v\delta\mathbf{U})_{i+1/2}^{m-} + \sum_{m=1}^2 \mathbf{k}_{i+1/2}^{m-} \Delta t \right) + \Delta t \mathbf{R}_i^n \end{aligned} \quad (39)$$

This is actually the basic Roe scheme formulated for the 1D shallow water equations. As it is well known, the basic explicit scheme requires some kind of correction called entropy fix, in order to avoid non-physical discontinuities near sonic points. As the proposed method is only an extension, the same entropy correction must be applied to the local advection velocities in those cases.

To extend the scheme to higher CFL the rule presented in (35) and (36) is used, and once the signals are distributed, the friction is computed pointwise and only once at cell i . The entropy correction, when necessary, is applied only once per interface before distributing the signals. The question of the applicability of the proposed scheme, based on a linear decomposition of the signals, remains open and will be evaluated in the section devoted to numerical results.

The numerical treatment at the boundaries when dealing with a non-linear system such as the 1D shallow water flow equations, follows closely the procedure described in Section 1.1.1 for the scalar case. In case of having a solid wall boundary condition no signal contributions are allowed to cross the boundary and are accumulated at the boundary cell, for conservation requirements where, afterwards, zero discharge is imposed. In case of dealing with a subcritical inlet or outlet boundary, one variable is externally imposed as physical boundary condition and the other variable is calculated using the updating information arriving to the boundary cell from the neighbours; some of the contributions cross the boundary as in Section 1.1.1. In case of having a supercritical outlet nothing is forced, the boundary cell is updated with the contributions from upstream neighbours and the rest of the contributions are allowed to cross the boundary.

1.5. One-dimensional test cases

1.5.1. An application to the linear wave equation. A numerical test case based on Equation (1) is first presented with $f = \lambda u$ using a uniform grid with 1000 cells, with a grid spacing $\Delta x = 1$, a constant value of $\lambda = 1$ and initial condition

$$u(x, 0) = \begin{cases} 2 & \text{if } 195 \leq x \leq 205 \\ 1 & \text{otherwise} \end{cases}$$

The simulation of the transport of this wave is performed using different values of CFL, equal to 0.95, 2.95, 4.95 and 9.95, according to the scheme proposed in Section 1.1. The results at time $t = 600$ s are shown on Figure 4, where it can be seen how the method becomes less diffusive as the parameter CFL increases, and the wave is transported along the domain in fewer steps. Table I displays the L_1 and L_∞ errors [16] computed comparing the numerical solutions and the exact solutions. At the same time, there is a simple and linear gain in the

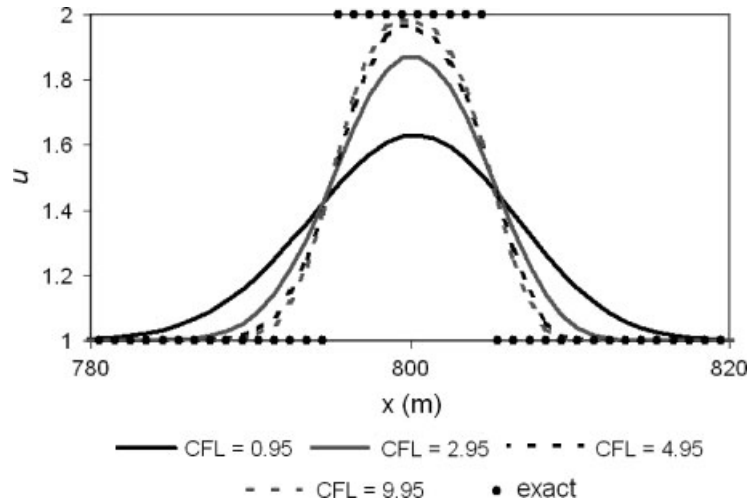


Figure 4. Linear advection of a square wave at $t = 600$ s for different values of CFL.

Table I. L_1 and L_∞ errors.

	CFL = 0.95	CFL = 2.95	CFL = 4.95	CFL = 9.95
L_1	8,550	5,192	3,756	3,266
L_∞	0,522	0,455	0,441	0,440

time step size as the CFL number is increased in linear problems on regular grids. Things are however different otherwise.

1.5.2. Application to 1D shallow water equations

1.5.2.1. Ideal dam break. The dam break flow subsequent to an initial 1m/5 m surface level discontinuity located at $x \leq 50$ m in a flat and frictionless domain 100 m long and 1 m wide is first presented. A uniform grid composed by 100 cells, $\Delta x = 1$ m is used and, following the scheme discussed in Section 1.5, different simulations are performed using different values of CFL. The boundaries act like solid walls, so that no flux can cross them. Figure 5 shows the comparison between the results of the water surface elevation $D = z + h$ and the exact solutions at different times $t = 3$ and 6 s (before interaction with the boundaries), for different values of CFL. At later times and when the value of the CFL increases over 8, the oscillations and instabilities linked to the effect of the accumulation of waves at the boundaries, makes the method fail. For values of $\text{CFL} < 5.0$ it can be seen that the differences in all cases are negligible. Table II shows the errors L_1 and L_∞ obtained comparing the exact and the numerical solution.

To illustrate how the instabilities are produced by the combination between high CFL values and the proposed accumulation technique at the closed boundaries, Figure 6 shows the results

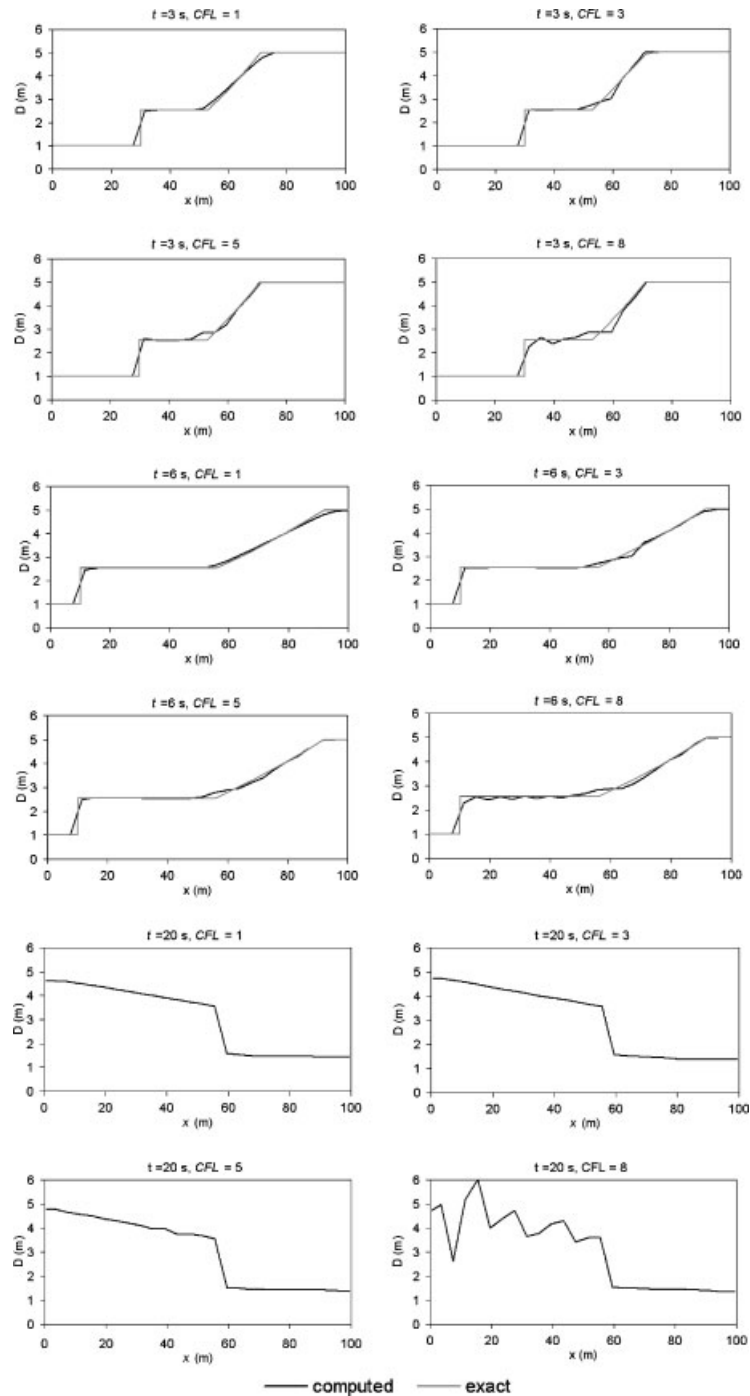


Figure 5. Dam break results for different values of CFL, NCELL = 100.

Table II. L_1 and L_∞ errors.

	CFL = 1	CFL = 3	CFL = 5	CFL = 8
$L_1, t = 3$ s	2,33	2,09	2,38	5,59
$L_\infty, t = 3$ s	0,21	0,26	0,29	0,41
$L_1, t = 6$ s	2,85	2,51	2,68	5,54
$L_\infty, t = 6$ s	0,17	0,20	0,22	0,24

when the total length is increased to 1000 m, using the same cell length and setting the initial dam break at $x = 500$ m. The solutions at $t = 30$ and 60 s are compared with the exact solutions. Again, when the value of the CFL increases over 8, the oscillations and instabilities produced due to the effect of the accumulation make the method fail. Table III shows the errors L_1 and L_∞ obtained comparing the analytical and the numerical solution. This is clearly a present day weakness of the method proposed and the best way to overcome it will be investigated in future work.

1.5.2.2. MacDonald steady-state open channel test case. MacDonald [17] supplied a set of realistic steady open channel flow test cases with analytical solution very well suited to validate the convergence of time-stepping schemes to steady state. The example used here consists of a rectangular channel of width $B = 10$ m and length $L = 150$ m, discretized in 150 cells. The Manning friction parameter is $n = 0.03$ (SI units) and the discharge is $Q = 20$ m³/s. According to MacDonald [17] the steady-state water depth, in meters, is for $0 \leq x \leq 100$:

$$h(x) = 0.741617 - \frac{0.25}{\tanh(3)} \tanh\left(3 \frac{(x-50)}{50}\right)$$

and for $100 < x \leq 150$

$$h(x) = \exp(-0.3(x-100)) \sum_{i=0}^4 k_i \left(\frac{(x-100)}{50}\right)^i + 1.7 \exp(0.005(x-150))$$

with $k_0 = -0.253363$, $k_1 = -1.18214$, $k_2 = 5.99444$, $k_3 = -118.907$ and $k_4 = 61.738$.

The bed slope analytical solution is

$$S_0(x) = \left(1 - \frac{Q^2}{9.08665(h(x))^3 B^2}\right) h'(x) + \frac{Q^2 n^2 (B + 2h(x))^{4/3}}{A^{10/3}}$$

and P and A in (38) become $P = B + 2h(x)$ and $A = Bh(x)$. This solution corresponds to a subcritical upstream inflow and a subcritical downstream outflow that, due to the bed variations, are connected by a smooth transition subcritical–supercritical followed by a discontinuous transition (steady hydraulic jump) supercritical–subcritical inside the domain. The scheme is applied starting from initial condition of uniform water depth, $h(x, t = 0) = 1$ m, and velocity $u(x, t = 0) = 2$ m/s, in all the domain. The upstream and downstream external boundary conditions are:

$$Q(x = 0, t) = 20 \text{ m}^3/\text{s}, \quad h(x = L, t) = 1.7 \text{ m}$$

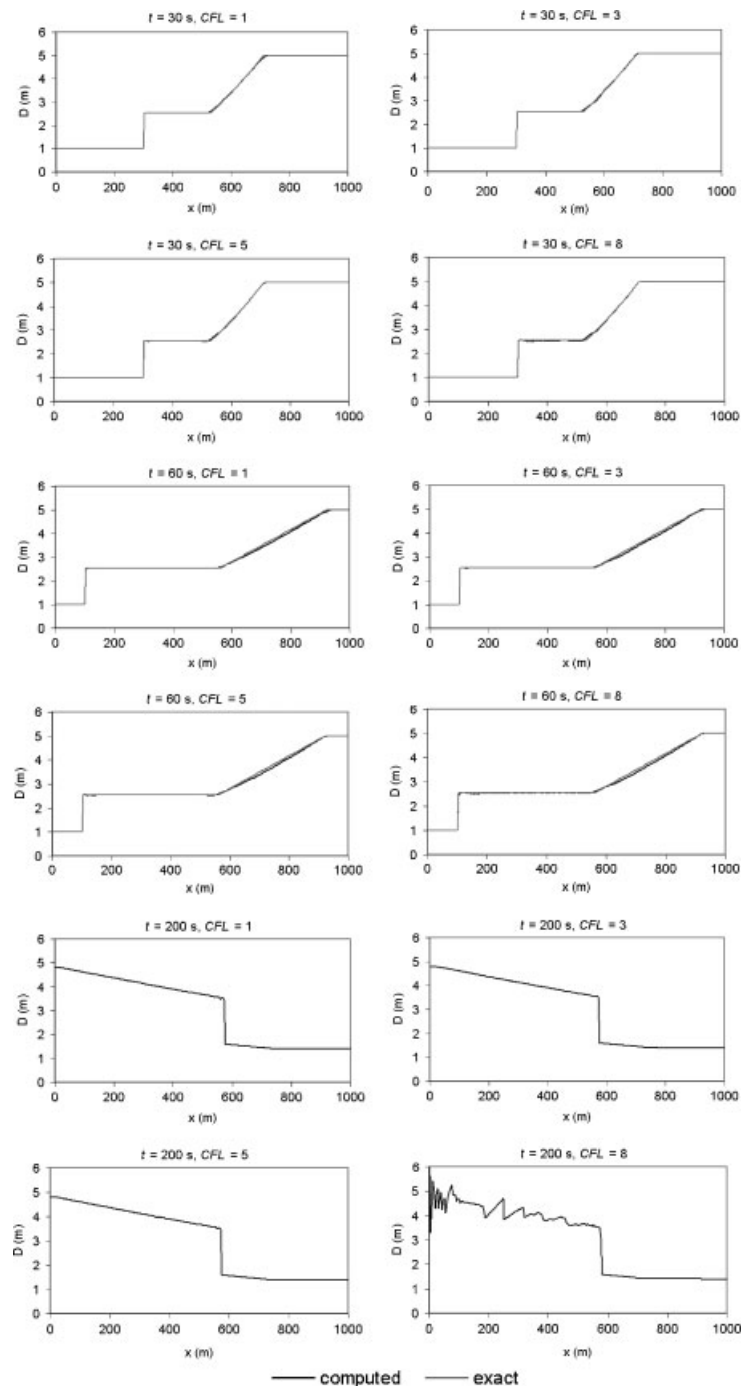
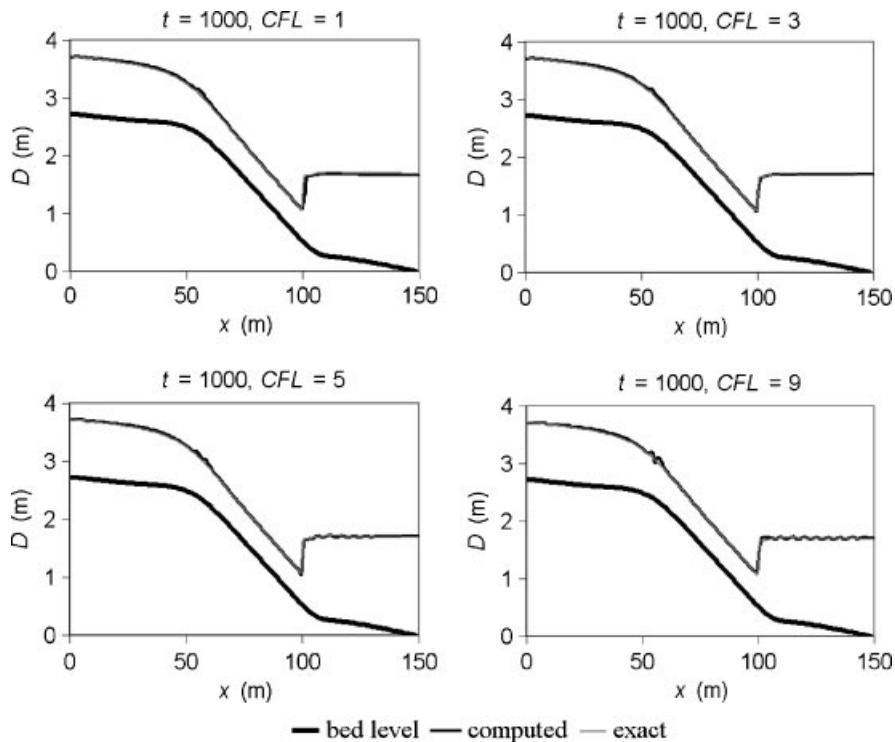


Figure 6. Dam break results for different values of CFL, NCELL = 1000.

Table III. L_1 and L_∞ errors.

	CFL = 1	CFL = 3	CFL = 5	CFL = 8
$L_1, t = 30$ s	5,20	4,45	3,72	8,30
$L_\infty, t = 30$ s	0,08	0,05	0,06	0,07
$L_1, t = 60$ s	24,63	25,30	25,71	29,97
$L_\infty, t = 60$ s	0,09	0,11	0,11	0,12

Figure 7. MacDonal's test case. Exact and numerical solution at $t = 1000$ s for CFL = 1, 3, 5 and 9.Table IV. MacDonal's test case. L_1 and L_∞ errors.

	CFL = 1	CFL = 3	CFL = 5	CFL = 9
L_1	3,344	2,149	2,066	2,449
L_∞	0,336	0,102	0,081	0,109

Figure 7 shows a plot of the numerical results for the water level surface, D , together with the channel bottom level for different values of CFL at $t = 1000$ s when the steady state has been reached. Table IV displays the numerical errors L_1 and L_∞ obtained comparing analytical and the numerical solutions.

2. NUMERICAL SCHEME. TWO-DIMENSIONAL CASE

2.1. Scalar equation

The 2D form of the scalar advection equation can be written as

$$\frac{\partial u}{\partial t} + \nabla \mathbf{f} = \frac{\partial u}{\partial t} + \boldsymbol{\lambda} \cdot \nabla u = 0, \quad \mathbf{f} = (f_x, f_y), \quad \boldsymbol{\lambda} = \frac{d\mathbf{f}}{du} \quad (40)$$

The upwind method described in previous sections will be generalized to this case. For that purpose the physical domain is discretized in triangular cells, as shown in Figure 8, and the function is represented using piecewise constant values u_i at the cells. The basic first order explicit upwind scheme applied to (40), with reference to Figure 8 gives the following rule for the updating of cell i :

$$A_i u_i^{n+1} = A_i u_i^n - \sum_{k=1}^3 (\tilde{\boldsymbol{\lambda}} \mathbf{n})_k^- (u_k - u_i) l_k \Delta t \quad (41)$$

where $(\tilde{\boldsymbol{\lambda}} \mathbf{n})_k^-$ represent the ingoing contributions from each surrounding edge k to cell i , A_i is the cell area, \mathbf{n}_k is the outward normal vector to the cell edge and l_k is the length of edge. Equation (41) can be rearranged to

$$u_i^{n+1} = u_i^n - \sum_{k=1}^3 \left(\Delta t \frac{l_k}{A_i} \tilde{\lambda}_k^- \right) \delta u_k = u_i^n - \sum_{k=1}^3 v_k \delta u_k \quad (42)$$

with $\tilde{\lambda}_k^- = (\tilde{\boldsymbol{\lambda}} \mathbf{n})_k^-$, $\delta u_k = u_k - u_i$ and $v_k = \Delta t \tilde{\lambda}_k^- (l_k / A_i)$. Also, the variable μ_k can be defined as $\mu_k = \text{int}(v_k)$. Note that μ_k is always a negative integer in the 2D model.

The classical CFL condition for explicit schemes on unstructured irregular grids states

$$\Delta t = \text{CFL} \Delta t_{\max}, \quad \text{CFL} \leq 1$$

with

$$\Delta t_{\max} = \min \{ \Delta t_{\max, i} \}_{i=1, \text{NCELL}} \quad (43)$$

$$\Delta t_{\max, i} = \min \left\{ \frac{A_{\min, k}}{|\boldsymbol{\lambda} \mathbf{n}|_k l_k} \right\}_{k=1, 3}, \quad A_{\min, k} = \min \{ A_k, A_i \}$$

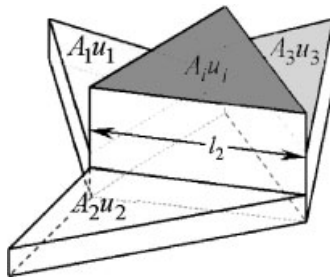


Figure 8. Domain discretization in triangles showing cell values and cell areas.

The updating waves at an edge k , can be rewritten as

$$v_k \delta u_k = \mu_k \delta u_k + (v_k - \mu_k) \delta u_k \quad (44)$$

or

$$|\mu_k \delta u_k + (v_k - \mu_k) \delta u_k| \leq N_k |\delta u_k| \quad (45)$$

with $N_k = (1 - \mu_k)$. Assuming a uniform mesh, in the 1D case, N_k gives the number of cells involved in the stencil, located in the direction given by $\tilde{\lambda}_k^-$, the first $N_k - 1$ cells are updated with δu_k and the contribution to cell N_k is $-(v_k - \mu_k) \delta u_k$. In the 2D case, the contribution at a cell edge when $|\mu_k| > 0$ must be spread among a number of cells given by

$$N_{2D,k} = \sum_{n=0}^{-\mu} 2^n \quad (46)$$

as explained below. It is worth signalling that this distribution must be performed in a genuinely 2D form, hence requiring unstructured triangular grids. As in the 1D irregular mesh case, when CFL values higher than one are sought, the parameters involved in the definition of Δt_{\max} must change involving a higher number of cells.

In order to introduce the procedure in a simple way, let us describe it with reference to Figure 9. When all the $\tilde{\lambda}_k^-$ at the edges surrounding cell i are evaluated and only $\tilde{\lambda}_k^- = \tilde{\lambda}_3^-$ is not null at the edge $(i, 3)$. If, in a first case $\text{CFL} \leq 1$ and $v_3 = 1$, the volume in cell i is increased with the quantity $\delta u_3 A_i$, and u_i^n is updated with δu_3 . Figure 9 shows the signal advance when $1 < \text{CFL} \leq 2$. This implies $\mu_k = -1$, and $N_{2D,k} = N_{2D,3} = 3$.

Following the previous philosophy, only the contribution $\delta u_3 A_i$ is assigned to i , and the quantity $-A_i (v_3 - \mu_3) \delta u_3$ is shared by cells 1 and 2. As the method must be conservative in

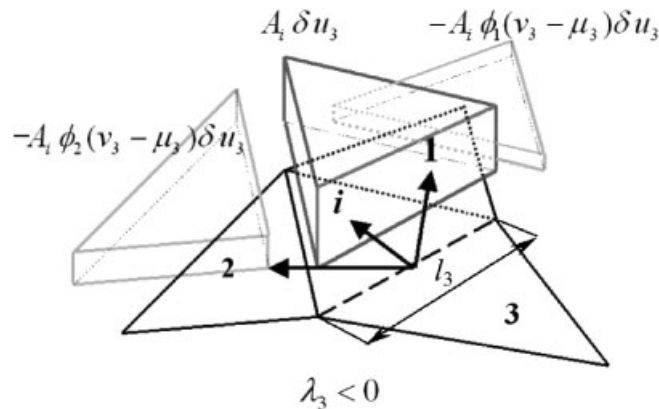


Figure 9. 2D Signal distribution. $1 < \text{CFL} \leq 2$.

each cell, the following splitting is proposed:

$$\begin{aligned} \delta u_3 & \text{ updates cell } i \\ -\phi_1 \frac{A_i}{A_1} (v_3 - \mu_3) \delta u_3 & \text{ updates cell 1} \\ -\phi_2 \frac{A_i}{A_2} (v_3 - \mu_3) \delta u_3 & \text{ updates cell 2} \end{aligned} \quad (47)$$

The proportion of the signal assigned to cells 1 and 2 depends on the area of the involved cells:

$$\phi_1 = \frac{A_1}{A_2 + A_1} < 1, \quad \phi_2 = \frac{A_1}{A_2 + A_1} < 1 \quad (48)$$

On the other hand, in order to develop a correct numerical method, it is necessary that

$$0 \leq \left| \phi_j \frac{A_i}{A_j} (v_3 - \mu_3) \right| \leq 1 \quad j = 1, 2 \quad (49)$$

as this factor controls the contributions to the involved cells, otherwise oscillations can appear. In a general situation, at every edge k of every cell i the following condition must be satisfied:

$$0 \leq \left| \frac{A_i}{A_j} (v_k - \mu_k) \right| \leq 1 \quad j = 1, 2 \quad (50)$$

Taking into account that $0 \leq |v_k - \mu_k| \leq 1$, the first option is to enforce $A_j \geq A_i$, obviously absurd. An alternative strategy is enforced by means of the search of $A_{\min,k}$ all over the cells involved in the stencil, and condition (43) has to be based on

$$A_{\min,k} = \min\{A_i, A_1, A_2, A_3\} \quad (51)$$

for the example considered.

Now the contributions are bounded by

$$\frac{A_i}{A_j} |v_k| = \frac{l_k |\lambda_k^-|}{A_j} \Delta t_{\max} \text{CFL} = \frac{A_{\min,k}}{A_j} \text{CFL} \leq \text{CFL} \quad (52)$$

and, as $\mu_k = \text{int}(v_k)$

$$\frac{A_i}{A_j} |v_k - \mu_k| = \frac{A_i}{A_j} |v_k| - \frac{A_i}{A_j} |\mu_k| \leq \text{CFL} - \text{int}(\text{CFL}) \leq 1 \quad (53)$$

Equation (50) is satisfied.

Furthermore, to smooth the size of the updating contributions in case of very irregular meshes, the following condition is imposed over the ϕ_j coefficients:

$$-\phi_j \frac{A_i}{A_j} (v_k - \mu_k) \leq \frac{1}{2}, \quad j = 1, 2 \quad (54)$$

As ϕ_j is a fixed shape factor (54) can be violated. The strategy in that case consists of relaxing the value of the CFL until (54) is granted. Using (51) and (54) the waves are not amplified and instabilities are avoided if CFL is kept less or equal to 2.

If the CFL value is chosen $2 < \text{CFL} \leq 3$, then $\mu_3 = -2$ and $N_{2D,3} = 7$. Figure 10 shows a sketch of the distribution of the contributions from cell edge $(i,3)$ in this case, with $\mu_3 = -2$. To ensure that the distribution is made with no amplification, the value of the time step must be calculated involving all the stencil cell areas, using $A_{\min,k}$ as

$$A_{\min,k} = \min\{A_i, A_1, A_{1,1}, A_{1,2}, A_2, A_{2,1}, A_{2,2}, A_3\} \tag{55}$$

In this case, following a conservative criterion again,

$$\begin{aligned} \delta u_3 & \text{ updates cell } i \\ (A_i/A_1)\phi_1\delta u_3 & \text{ updates cell } 1 \\ -(A_i/A_{1,1})\phi_1\phi_{1,1}(v_3 - \mu_3)\delta u_3 & \text{ updates cell } 1, 1 \\ -(A_i/A_{1,2})\phi_1\phi_{1,2}(v_3 - \mu_3)\delta u_3 & \text{ updates cell } 1, 2 \\ (A_i/A_2)\phi_2\delta u_3 & \text{ updates cell } 2 \\ -(A_i/A_{2,1})\phi_2\phi_{2,1}(v_3 - \mu_3)\delta u_3 & \text{ updates cell } 2, 1 \\ -(A_i/A_{2,2})\phi_2\phi_{2,2}(v_3 - \mu_3)\delta u_3 & \text{ updates cell } 2, 2 \end{aligned} \tag{56}$$

where, for instance,

$$\phi_{1,1} = \frac{A_{1,1}}{A_{1,1} + A_{1,2}} < 1, \quad \phi_{1,2} = \frac{A_{1,2}}{A_{1,1} + A_{1,2}} < 1 \tag{57}$$

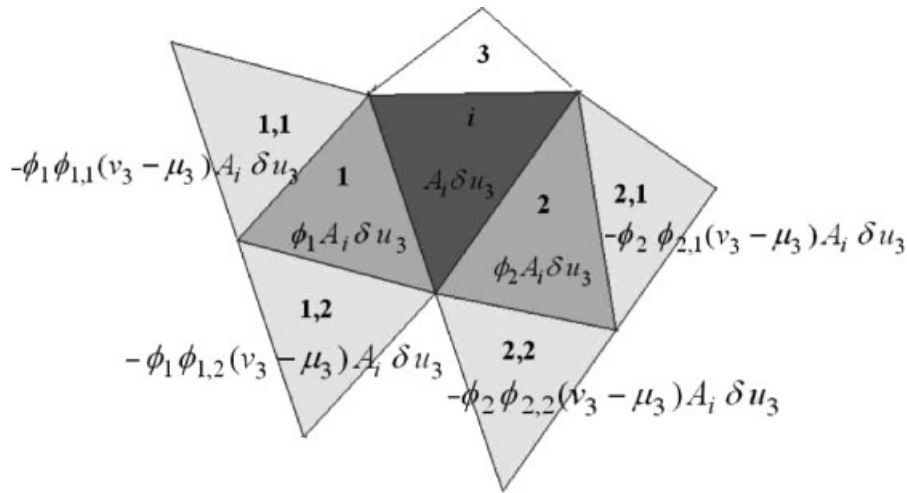


Figure 10. Signal volume distribution. $2 < \text{CFL} \leq 3$.

In order to avoid again too large contributions arising from very different cell sizes, the following condition is set for the four most distant cells from cell i :

$$-\frac{A_i}{A_{j,l}} \phi_j \phi_{j,l} (v_k - \mu_k) \leq \frac{1}{4}, \quad j=1,2 \quad \text{and} \quad l=1,2 \quad (58)$$

If (58) is not satisfied, the CFL value is reduced.

For a general value of CFL, the time step must be defined based on the maximum number of involved cells

$$A_{\min,k} = \min\{A_i, A_{j=1, N_{2D,k}}\} \quad (59)$$

and, according to the previously detailed philosophy, the following condition is imposed over the $2^{-\mu_k}$ distribution coefficients of the signals reaching the q farthest cells:

$$-\frac{A_i}{A_{j,l,\dots,q}} \phi_j \phi_{j,l} \dots \phi_{j,l,\dots,q} (v_k - \mu_k) \leq 2^{\mu_k}, \quad j=1,2, \quad l=1,2 \dots q=1,2 \quad (60)$$

However, as will be seen in the following sections, in some cases condition (60) is not sufficient to keep the solution free from oscillations. As the proposed scheme is highly geometry dependent, when the cell size varies strongly from cell to cell, condition (60) may not be enough. To ensure a well-behaved solution, condition (60) must be extended to the sets of cells that share the same contributions. Therefore, the following inequalities must be simultaneously checked over the j, l, \dots, q cells:

$$\begin{aligned} \frac{A_i}{A_j} \phi_j (v_k - \mu_k) &\leq \frac{1}{2}, \quad j=1,2 \\ \frac{A_i}{A_{j,l}} \phi_j \phi_{j,l} (v_k - \mu_k) &\leq \frac{1}{4}, \quad j=1,2, \quad l=1,2 \\ &\vdots \\ -\frac{A_i}{A_{j,l,\dots,q}} \phi_j \phi_{j,l} \dots \phi_{j,l,\dots,q} (v_k - \mu_k) &\leq 2^{\mu_k}, \quad j=1,2, \quad l=1,2 \dots q=1,2 \end{aligned} \quad (61)$$

When, for a target CFL value, one inequality of the set in (61) is violated, the CFL value must be reduced, leading automatically to a reduction in the number of involved cells, and in consequence, to a reduction in the number of involved conditions.

2.2. Boundary conditions treatment

As in the 1D problem, in the case of open boundaries nothing special has to be done. When the boundaries are closed or solid walls, the technique mentioned in Section 1.2 is also adapted to 2D meshes. In this case, the boundary cell may have one or two edges acting like solid walls (Figure 11). When two edges act like solid walls all the contributions are added to the same cell (left part of Figure 11). If the boundary cell has a closed wall boundary edge and an open boundary edge, some of the contributions are let to go out through the open boundary (right part of Figure 11).

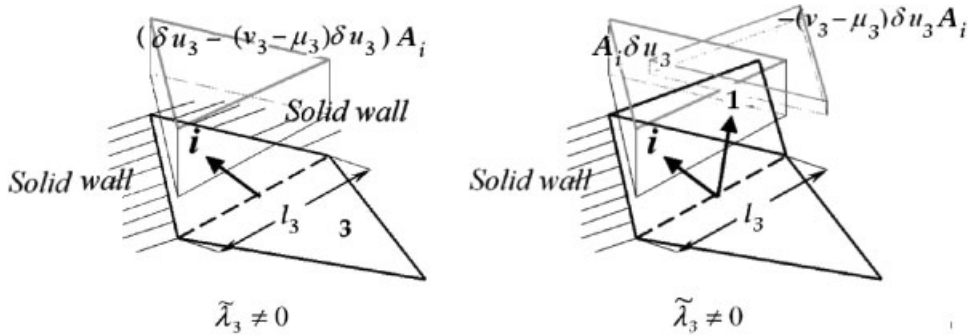


Figure 11. Solid wall boundary conditions, with $1 < CFL \leq 2$.

2.3. Scalar equation with source terms

When a source term $S = S(u, x, y)$ is added to (40)

$$\frac{\partial u}{\partial t} + \lambda \cdot \nabla u = S(u, x, y) \tag{62}$$

the resulting equation can be solved via an extension of the method outlined in Section 1.3:

$$u_i^{n+1} = u_i^n - \sum_{k=1}^3 v_k \delta u_k + \sum_{k=1}^3 S_k^- \frac{l_k}{A_i} \tag{63}$$

with $S_k^- = \frac{1}{2} S_k (\tilde{\lambda}_k - |\tilde{\lambda}_k|)$ and S_k a suitable discretization of the source term at cell edge k . This equation can be transformed into

$$u_i^{n+1} = u_i^n - \sum_{k=1}^{NE} (v_k \delta u_k + \kappa_k \Delta t) \tag{64}$$

with $\kappa_k = -S_k^- (l_k/A_i)$.

The scheme proposed here to propagate the signals, including the source term, from the cell edges is the same as explained in Section 2.1 for the pure advection but corrected in the way described for the 1D case with source terms. If, for instance, $1 < CFL \leq 2$ as in Figure 9, cell i would be updated with the quantity

$$(\delta u_3 + (\kappa \Delta t)_3) \tag{65}$$

and cells 1 and 2 would receive, respectively,

$$-\frac{A_i}{A_1} \phi_1 (v_3 \delta u_3 + (\kappa \Delta t)_3), \quad -\frac{A_i}{A_2} \phi_2 (v_3 \delta u_3 + (\kappa \Delta t)_3) \tag{66}$$

being $\Delta t_{2,k} = -(v_k - \mu_k) \Delta t_{1,k}$ and $\Delta t_{1,k} = -(A_i/\tilde{\lambda}_k^- l_k)$.

For bigger CFL values, the signals are transported taking into account the criteria and conditions presented in the previous sections. Finally, it can be checked that, when $\mu_k = 0$ automatically $\Delta t_{2,k} = \Delta t_{\max}$.

2.4. Application to a 2D system of equations with source terms

As in Section 1.4, it is assumed that Roe's linearization can be used to decouple the hyperbolic system

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x}(\mathbf{U}) + \frac{\partial \mathbf{G}}{\partial y}(\mathbf{U}) = \mathbf{S}(x, y, \mathbf{U}) \quad (67)$$

by means of the normal flux Jacobian matrix

$$\mathbf{J}_n = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} n_x + \frac{\partial \mathbf{G}}{\partial \mathbf{U}} n_y \quad (68)$$

making possible the generation of an approximate matrix \mathbf{J}_n^* , whose eigenvalues $\tilde{\lambda}^m$ and eigenvectors $\tilde{\mathbf{e}}^m$ [2, 4], can be used to define the signals, and also that an upwind treatment can be done over the source terms. In this form (67) can be discretized following

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{k=1}^{\text{NE}} \sum_m ((\tilde{\lambda}^{m-} \alpha^m - \beta^m) \tilde{\mathbf{e}}^m)_k^n \frac{I_k}{A_i} \Delta t \quad (69)$$

which can also be written as

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \sum_{k=1}^{\text{NE}} \sum_m (v^m \delta \mathbf{U}^m + \boldsymbol{\kappa}^m \Delta t)_k^n \quad (70)$$

where

$$\boldsymbol{\kappa}^m = \frac{I_k}{A_i} \beta^m \tilde{\mathbf{e}}^m \quad \text{and} \quad v^m = \lambda^{m-} \frac{I_k}{A_i} \Delta t$$

The value of $1 - \mu_k^m$, with $\mu_k^m = \text{int}(v_k^m)$ indicates the number of cells reached by the contributions generated by each component of $\delta \mathbf{U}_k$, $\delta \mathbf{U}_k^m$ at each cell edge, according to the method described in the previous sections. In this case, the search of the global parameter Δt_{\max} involves all the eigenvalues, so that for each value of CFL the value of the time step must be recalculated as

$$\Delta t = \text{CFL} \Delta t_{\max}$$

with

$$\begin{aligned} \Delta t_{\max} &= \min \{ \Delta t_{\max, i} \}_{i=1, \text{NCELL}} \\ \Delta t_{\max, i} &= \min \left\{ \frac{A_{\min, k}}{|\lambda \mathbf{n}|_k I_k} \right\}_{k=1, 3}, \quad A_{\min, k} = \min \{ A_i, A_{j=1, \text{N}_{2\text{D}, k}} \} \end{aligned} \quad (71)$$

Again, taking in account (60), the next condition must be ensured for each m wave over the $2^{-\mu_k}$ distribution coefficients of the signals reaching the q farthest cells:

$$-\frac{A_i}{A_{j, l, \dots, q}} \phi_j \phi_{j, l} \dots \phi_{j, l, \dots, q} (v_k^m - \mu_k^m) \leq 2^{\mu_k^m}, \quad j = 1, 2, \quad l = 1, 2 \dots q = 1, 2 \quad (72)$$

or, following the most strict criteria in (61)

$$\begin{aligned}
 \frac{A_i}{A_j} \phi_j (v_k^m - \mu_k^m) &\leq \frac{1}{2}, \quad j = 1, 2 \\
 \frac{A_i}{A_{j,l}} \phi_j \phi_{j,l} (v_k^m - \mu_k^m) &\leq \frac{1}{4}, \quad j = 1, 2, \quad l = 1, 2 \\
 &\vdots \\
 -\frac{A_i}{A_{j,l,\dots,q}} \phi_j \phi_{j,l} \dots \phi_{j,l,\dots,q} (v_k^m - \mu_k^m) &\leq 2\mu_k^m, \quad j = 1, 2, \quad l = 1, 2 \dots q = 1, 2
 \end{aligned} \tag{73}$$

For both criteria, if for a target CFL value one inequality is violated, the time step must be reduced.

The source terms are involved through the coefficients

$$\Delta t_{2,k}^m = -(v_k^m - \mu_k^m) \Delta t_{1,k}^m, \quad \Delta t_{1,k}^m = -\frac{A_i}{\tilde{\lambda}_k^- l_k} \tag{74}$$

according to the method previously described. It can be checked that, when $\mu_k^m = 0$, automatically $\Delta t_{2,k}^m = \Delta t_{\max}$. Bigger values of CFL involve more cells, and each component of $\delta \mathbf{U}_k$ must have an analogous treatment as explained in previous sections.

2.5. Influence of the mesh

The simulation of 2D problems using techniques based on finite difference methods or structured finite volumes has a clear dependence on the mesh used. The domain is divided into cells and the updating waves are computed following the normal vector to their edges and depending on the cells size. In recent times, the introduction of unstructured meshes has reduced this influence, by means of directionless elements, in special those created by Delaunay solvers [18]. The influence of the mesh in the solution is, however, more evident when values of $\text{CFL} > 1$ are used in the context of a method such as the one presented in this work, and two essential factors are found: the direction of the normal vector to the cell edge and the cell area.

In practice, it is impossible to make the waves advance farther than the adjacent cell in structured meshes, as the wave direction, completely dominated by the geometry mesh, is governed by the x and y aligned edges. This tendency is also noticeable on unstructured meshes generated by means of frontal-advancing solvers, where the advance can generate cells following the main directions. As the numerical contributions between cells are computed following the normal to the edges, the equations are forced to give solutions depending on the direction of those normals, that is, aligned in some way with the main axis, avoiding the possibility of transporting the waves in other directions. In meshes generated by means of a Delaunay solver it is possible the use of values of $\text{CFL} > 1$ since, even though local main directions can appear, they can be avoided by means of retriangularization. The retriangularization can be simply done by adding nodes at the geometrical centre of each cell. Furthermore, another geometric parameter has to be taken into account: the uniformity in the cell size. Conditions (60) and (61) lead to maximum values of CFL when the cell areas are of uniform size

and lead to distribution coefficients ϕ approximately equal to 1/2. When local mesh refinement is required, the quality of the mesh is given by the maximum smoothness in the cell size variation.

2.6. Application to 2D shallow water equations

The 2D shallow water equations, which represent vertically averaged mass and momentum conservation, form a system like (67) with:

$$\mathbf{U} = (h, q_x, q_y)^T$$

$$\mathbf{F} = \left(q_x, \frac{q_x^2}{h} + \frac{gh^2}{2}, \frac{q_x q_y}{h} \right)^T, \quad \mathbf{G} = \left(q_y, \frac{q_x q_y}{h}, \frac{q_y^2}{h} + \frac{gh^2}{2} \right)^T \quad (75)$$

where $q_x = uh$ and $q_y = vh$. The variable h represents the water depth, g is the acceleration of the gravity and (u, v) are the averaged components of the velocity vector \mathbf{u} along the x and y coordinates, respectively. The source terms in the momentum equations are the bed slopes and the friction losses along the two coordinate directions,

$$\mathbf{S} = (0, gh(S_{0x} - S_{fx}), gh(S_{0y} - S_{fy}))^T \quad (76)$$

where

$$S_{0x} = -\frac{\partial z}{\partial x}, \quad S_{0y} = -\frac{\partial z}{\partial y}$$

and the friction losses in terms of the Manning's roughness coefficient, with

$$S_{fx} = n^2 u \sqrt{u^2 + v^2} / h^{4/3}, \quad S_{fy} = n^2 v \sqrt{u^2 + v^2} / h^{4/3}$$

The system has a normal flux Jacobian matrix, \mathbf{J}_n , whose eigenvalues are a representation of the characteristic speeds. The details and specific form of the involved variables can be found for instance in References [2, 4].

The source term vector can also be decomposed in two different parts that will be treated separately: those based on spatial derivatives, such as the bottom variations \mathbf{B} and the rest, in our case the friction term \mathbf{R} , $\mathbf{S} = \mathbf{B} + \mathbf{R}$. 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-state cases. The friction term \mathbf{R} is discretized in a pointwise manner, so that the final expression for the numerical scheme is

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{k=1}^{\text{NE}} \sum_{m=1}^3 ((\tilde{\lambda}^{m-} \boldsymbol{\alpha}^m - \beta^{m-}) \tilde{\mathbf{e}}^m)_k \frac{ds_k}{A_i} \Delta t + \Delta t (\mathbf{R})_i^n \quad (77)$$

This expression is a rule to update the value of the conserved variables in each cell using the information only from the surrounding cell edges, when the value of the parameter CFL is less or equal one. When this parameter is increased, the technique to follow is detailed in the previous sections, taking into account that the CFL applied produces variations in the value of Δt_{\max} .

2.7. Two-dimensional test cases

2.7.1. Application to the 2D linear advection equation. A periodic wave is transported through a square domain defined by $0 \leq x \leq 1$ and $0 \leq y \leq 1$. A constant and uniform diagonal advection velocity $\lambda = (1, 1)^T$ is used to advect the initial condition defined by

$$u(x, y, 0) = \cos(2\pi x) \cos(2\pi y)$$

Periodic boundary conditions are imposed to ensure that the initial solution, (Figure 12, left), and the solution after every second are equal, as the periodic wave returns to its initial position. The domain was discretized by a Delaunay mesh solver (BAMG) in 2024 cells (Figure 12, right).

The data in the elements are initialized to the average value of the exact solution over each triangle and the errors are computed by comparing the average exact solution with the numerical piecewise constant values at each cell after the first period. Table V shows the CFL attained when a target CFL value is set and (60) or (61) is enforced. It also displays how, in general, the error decreases to a limit as the CFL increases. Figure 13 displays the contour plot solutions after one period, for different values of CFL, and how the peak amplitude is better preserved when the CFL grows.

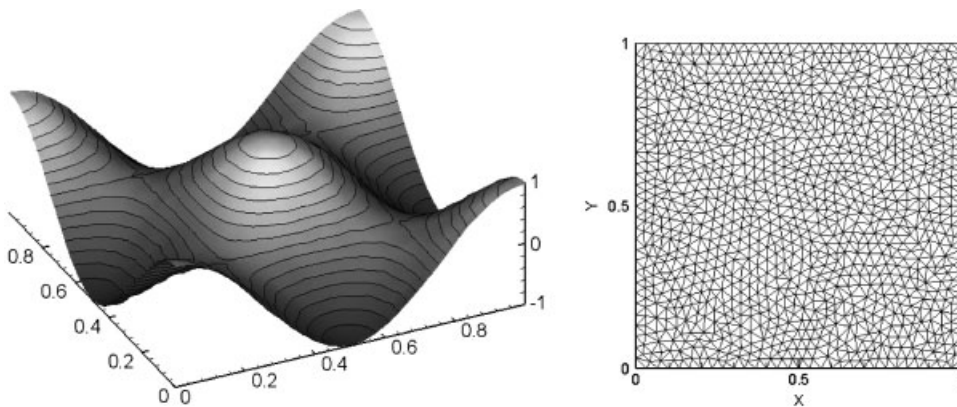


Figure 12. Initial and exact solution every second (left) and mesh discretization (right).

Table V. Maximum CFL attained depending on the target CFL.

Target CFL	Max. CFL (60)	Max. CFL (61)	L_1 (60)	L_1 (61)	L_∞ (60)	L_∞ (61)
1	1	1	0,049	0,049	0,321	0,321
2	2	2	0,040	0,040	0,222	0,222
3	2.572	2.572	0,039	0,039	0,177	0,177
4	2.653	2.572	0,038	0,039	0,182	0,177

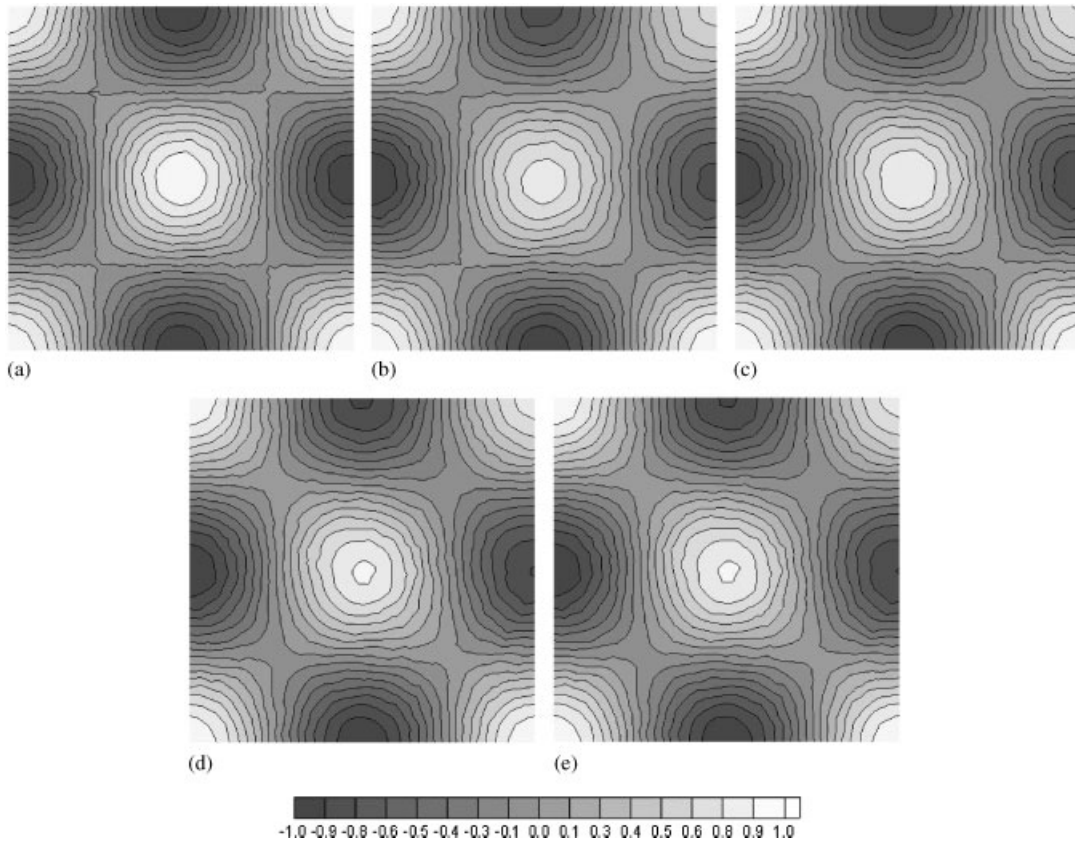


Figure 13. Exact solution (a) and numerical solution using CFL = 1 (b), CFL = 2 (c), CFL = 3 (d) and CFL = 4 (e), after one period using condition (60).

2.7.2. Application to the 2D shallow water equations

2.7.2.1. *Steady-state test case.* A 2D steady flow test case with analytical solution is used to validate the convergence of the proposed scheme. The flow discharge is constant in all the domain and equal to

$$q_x(x, y) = q_x, \quad q_y(x, y) = q_y$$

and the steady-state water depth and bed slope analytical functions are

$$h(x, y) = a + q_x x + q_y y, \quad z(x, y) = -\frac{1}{2g} \frac{(q_x^2 + q_y^2) + 2g(a + q_x x + q_y y)^3}{(a + q_x x + q_y y)^2}$$

The performance of the scheme is tested using $q_x = 0.3$, $q_y = 0.4$ and $a = 1$, in a squared domain, defined by $0 \leq x \leq 10$ and $0 \leq y \leq 10$. The assumed values correspond to an almost diagonal subcritical flow over a frictionless bottom of variable bed slope in which two sides

of the domain are inlet open boundaries and the other two sides are outlet open boundaries. This domain was discretized using the same Delaunay mesh solver (BAMG). To ensure that no privileged directions were presented, the initial mesh was retriangularized leading to 4602 cells. The steady water level is computed starting from initial condition of still water:

$$\begin{aligned} h(x, y, 0) + z(x, y, 0) &= 0 \\ u(x, y, 0) &= 0, \quad v(x, y, 0) = 0 \end{aligned}$$

To achieve the desired solution the unit discharges are the conditions imposed at the boundaries of the domain (subcritical flow). Figure 14 shows a 3D view of the exact solution of the water surface level (a), the contour plot of the bottom level (b) and a detail of the mesh discretization (c).

The error in the water depth is computed comparing the average exact solution of the water depth with the numerical piecewise constant values defined in each cell when the steady state has been achieved at time $t = 200$ s. Table VI shows the L_1 error and the L_∞ error *versus* the target CFL value. The results obtained using (72) and (73) are the same for target CFL values equal to 1, 2 and 3. When (73) is enforced, the results for values of $\text{CFL} > 3$ are equal. No values are displayed for values of $\text{CFL} > 3$ for condition (72), as in this example (72) is not strong enough to guarantee stability and the performance fails.

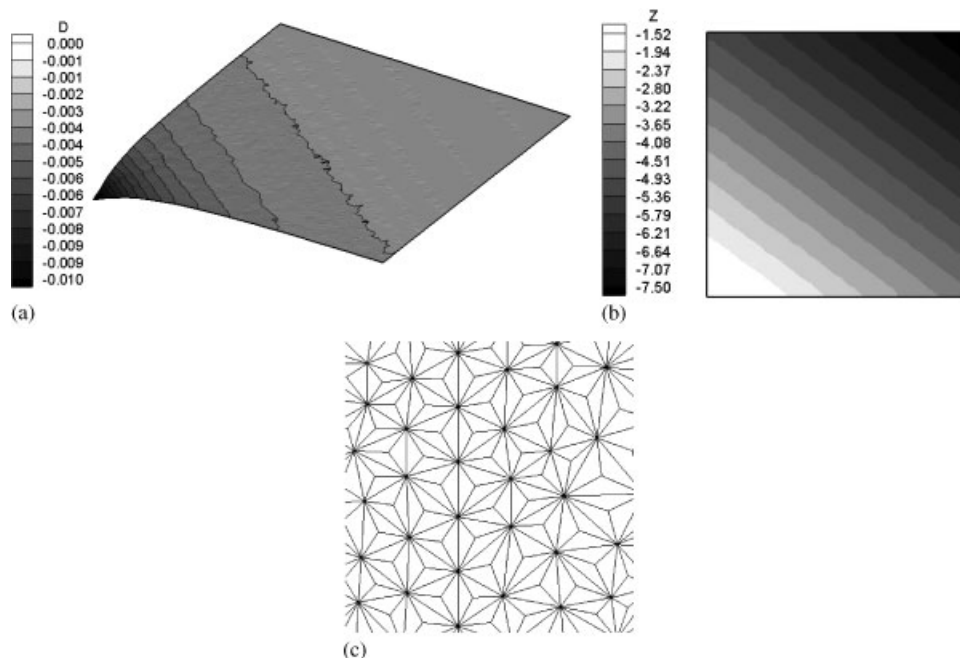


Figure 14. (a) 3D view of the steady water surface level; (b) bottom level; and (c) mesh discretization detail.

Table VI. Evolution of the L_1 and L_∞ errors for various values of CFL.

Target CFL	L_1 (72)	L_1 (73)	L_∞ (72)	L_∞ (73)
1	0,0972859	0,0972859	0,001292	0,001292
2	0,0972859	0,0972859	0,001292	0,001292
3	0,0972859	0,0972859	0,001292	0,001292
4	—	0,0972859	—	0,001292

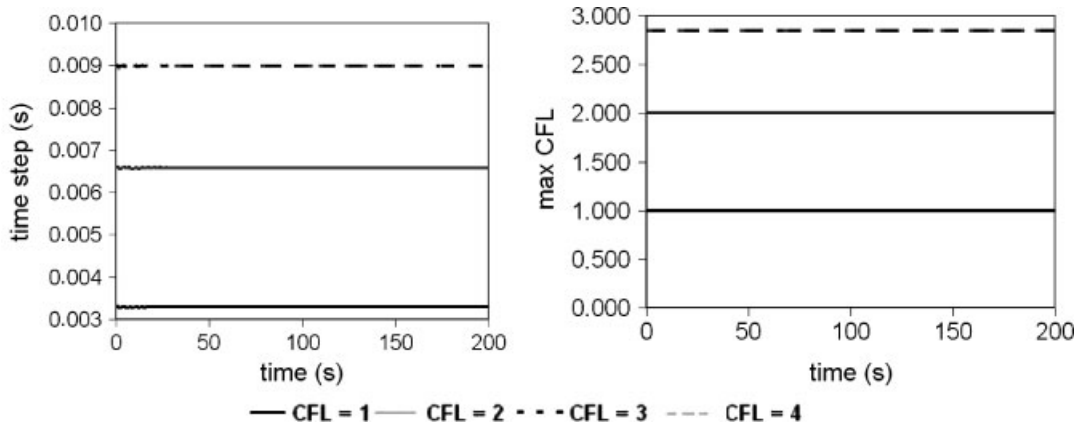


Figure 15. Evolution of the time step size in time (left) and evolution of the maximum allowable CFL using (73).

The magnitude of the time step and the maximum CFL value achieved are displayed in Figure 15 for target values of $\text{CFL}=1, 2, 3$ and 4 when (73) is enforced. In this case the results for values of $\text{CFL}=3$ and 4 are identical due to the CFL reduction induced by the grid-dependent condition (73); bigger values of the target CFL parameter lead to the same limitation in this example.

Figure 16 shows the contour plots of the water depth for the exact solution using the proposed scheme with $\text{CFL}=3$ and condition (73). Despite the different magnitude in the time step when using the method for various CFL values, no differences are found between the numerical solutions.

2.7.2.2. Asymmetric dam break in a laboratory model. The next validation test case for the proposed model deals with a 2D dam break flow problem. The experiment was carried out at the CITEEC, Coruña, Spain [19]. The set-up consists of a closed pool divided in two parts (Figure 17, left) by a removable gate. In this case the bed is plain. The experiment was performed for an initial depth ratio of $0.5/0.1$ m assuming a Manning roughness parameter $n=0.01$. The initial mesh was generated using a Delaunay solver (Triangle) and retriangular-

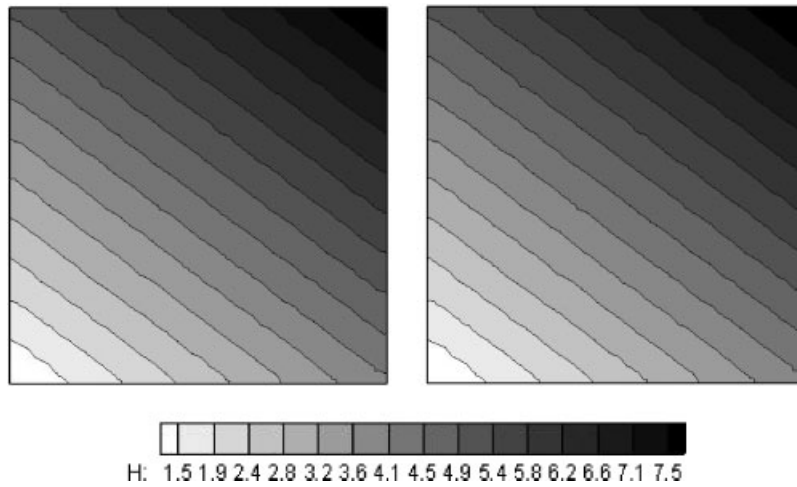


Figure 16. Exact water depth in meters (left) and water depth computed using $CFL=3$ at $t=200$ s (right) and condition (73).

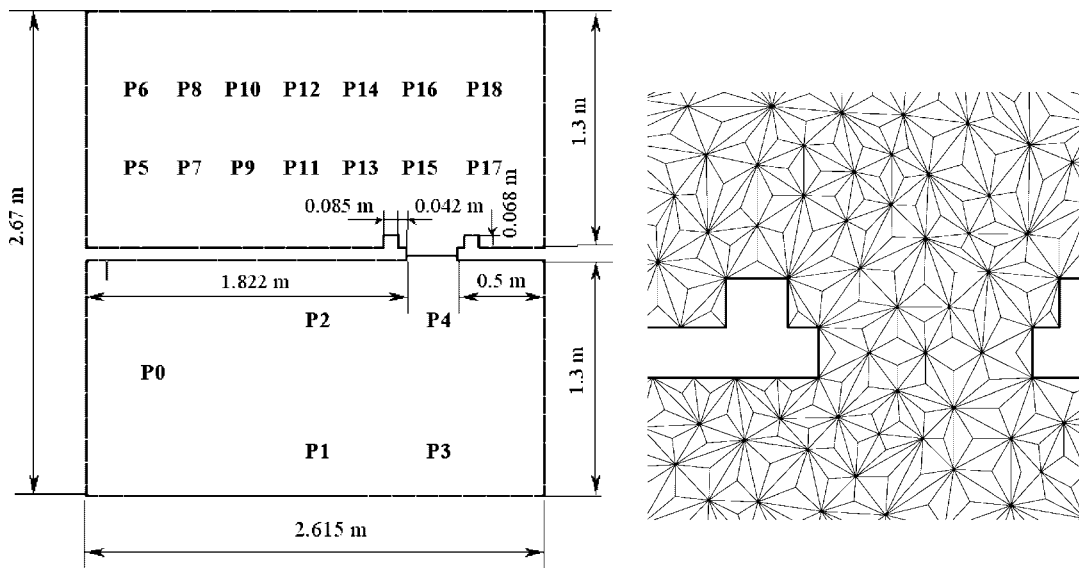


Figure 17. Model geometry (left) and detail of the mesh (right).

ized as indicated above leading to 7875 cells (Figure 17, right). The results are compared with the measures provided by 17 gauge points distributed all over the domain (Figure 17, left).

Figure 18 (left) shows the variation of the time step in function of time, for target values of CFL 1, 2, 3 and 4. In this case, condition (72) is sufficient, and conditions (72) and (73) lead to the same result. Figure 18 (right) also displays the maximum CFL allowed for various values of target CFL. As can be seen from the figure, the target value of CFL = 4 is the first to be affected by the grid-dependent restrictions imposed by condition (72). Values of the target parameter above 4 would lead to the same actual time step size.

Figure 19 shows contour level maps of water level at different times using target values of CFL 1 and 4. Figure 20 shows 3D plots of the water surface at time 2 s for both values of CFL. As it can be seen, the differences between them are negligible. The same conclusion is achieved about the vector velocity maps shown in Figure 21. The comparison between numerical results and experimental data at some gauging points of the physical model is displayed in Figure 22, where the numerical results for the different values of CFL are coincident in a single line.

2.7.2.3. Dam break in a constricted channel. Next application deals again with a dam break problem in a channel with a severe constriction that notably affects the propagation of the wave. The construction and the experimental measures in this test case were developed in the Laboratorio Nacional de Ingenieria Civil, Lisbon (Portugal) [20]. The channel presents no slope, and the walls are high enough to avoid the flow from crossing the boundaries (Figure 23). The wave advance produced by the dam break is reflected partially by the narrowing, and the wave suffers attenuation in its downstream advance. The roughness parameter n assumed is 0.01 (SI units). Again, an initial mesh is generated using a Delaunay solver (Triangle) and, to ensure that no privileged directions are present, is retriangularized. The final computational mesh contains 42 621 cells. The results are compared with the measures provided by 4 gauges distributed along the axis of the domain. A removable gate is located at $x = 6.1$ m, the initial water level is 0.3 m upstream of the gate and 0.003 m downstream of it.

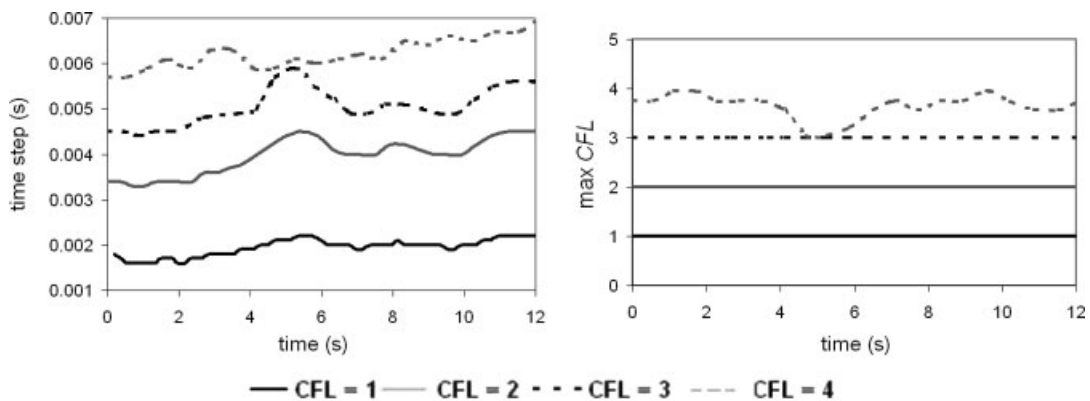


Figure 18. Evolution of the time step (left) and the maximum allowable CFL (right), using (72) or (73).

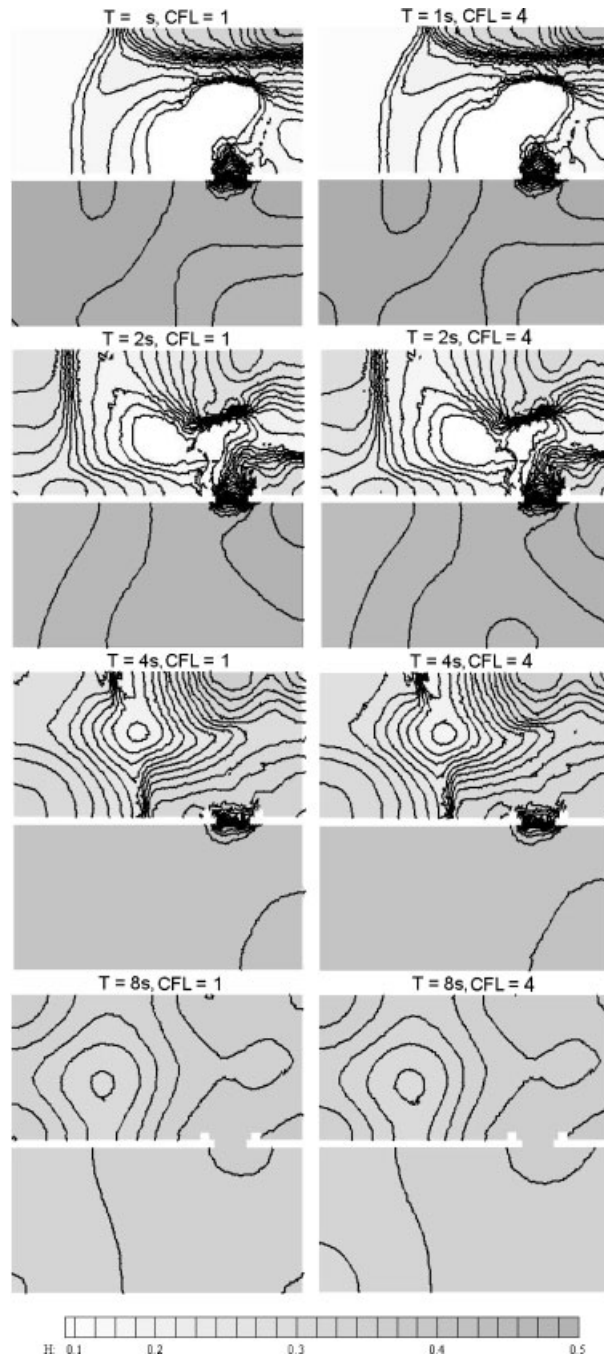


Figure 19. Water level, in meters, at times $t = 1, 2, 4$ and 8 s, for values of CFL = 1 (left column) and 4 (right column).

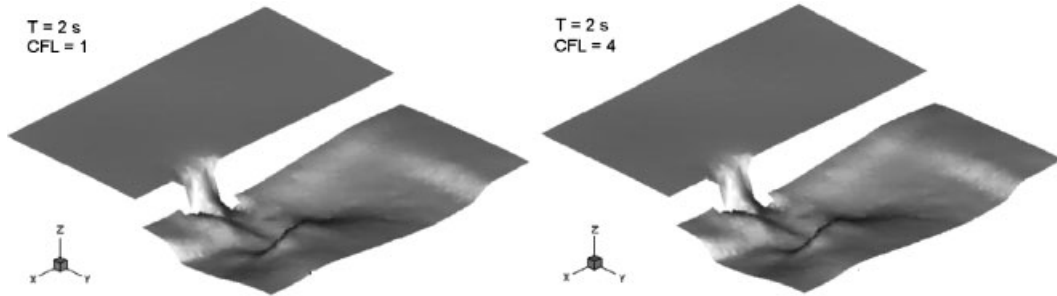


Figure 20. Surface water levels at time 2 s, computed with CFL = 1 and 4.

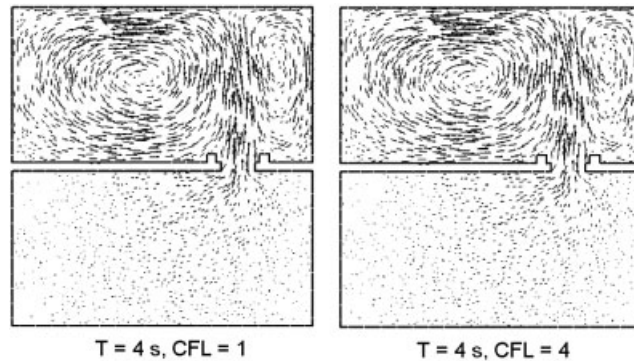


Figure 21. Velocity vectors at time 4 s, computed with CFL = 1 (left) and 4 (right).

Figure 24 shows the variation of the time step and the variation of the maximum allowed CFL in function of time, for target values of CFL 1, 2, 3, and 4. The previous tendencies are found in this new case, and both conditions (72) and (73) produce the same variations in the time step.

The iso-level maps shown in Figure 25 and the 3D plots shown in Figure 26 present almost imperceptible differences in the results for CFL 1 or 4. Also, the vector velocity maps displayed in Figure 27 are almost identical despite the variation of the parameter CFL. The results for the different values of CFL are compared with the measures provided by the 4 gauges in Figure 28. Although the numerical solution is not able to reproduce all the details of the experimental evolution in water depth, this is due to a limitation in the shallow water applicability to this case and not to the numerical technique. The results show almost imperceptible differences for all the CFL values tested.

2.7.2.4. Variable bottom in the constricted channel dam-break. To see how the proposed scheme deals with highly variable bed slope in presence of transient flow a last numerical experiment is performed using the same mesh and geometry of the previous case with a

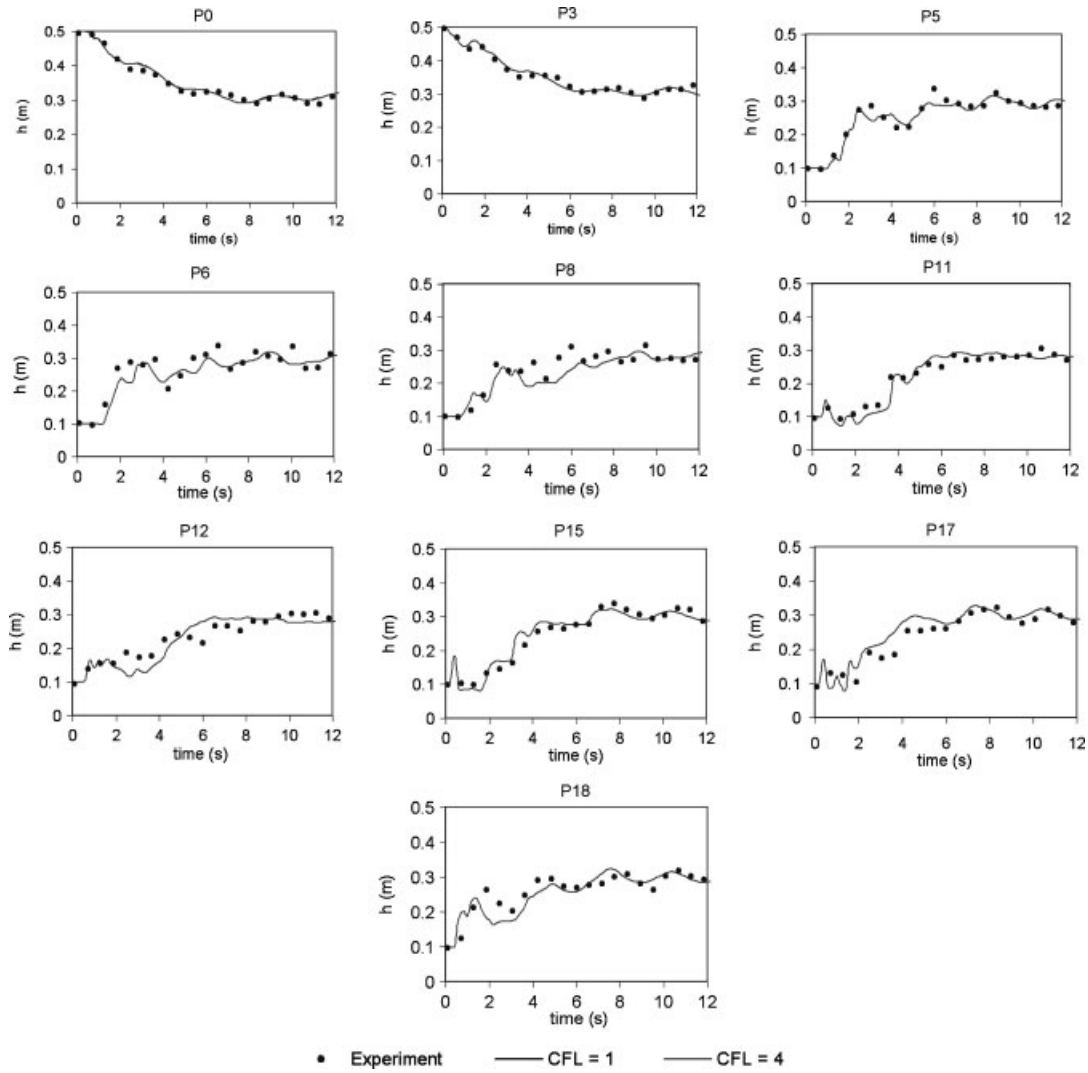


Figure 22. Comparison between experimental data and numerical data for CFL = 1 and 4.

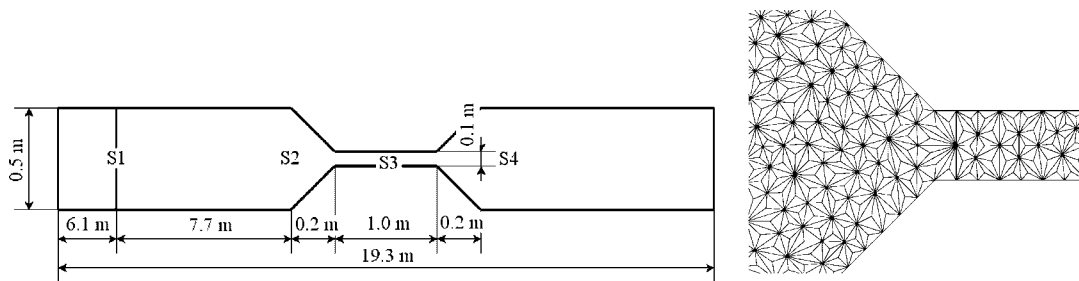


Figure 23. Model geometry (left) and mesh detail (right).

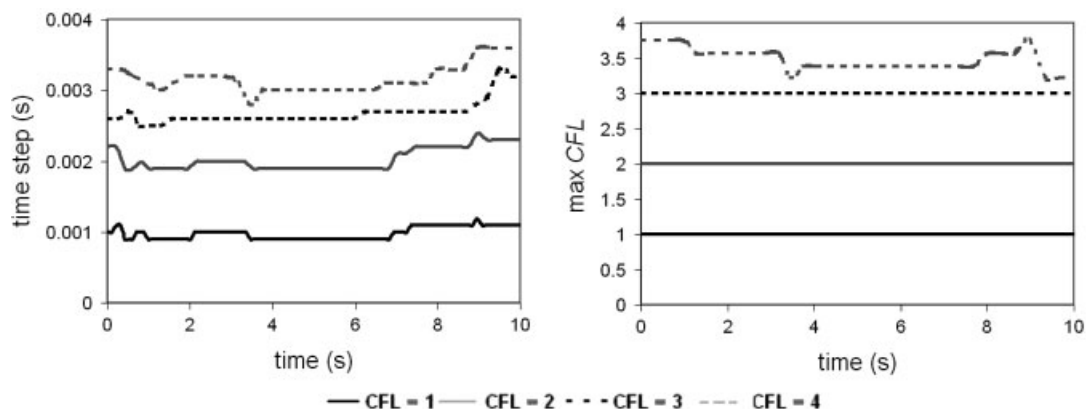


Figure 24. Evolution of the time step (left) and maximum allowable CFL (right) in function of time using (72) or (73).

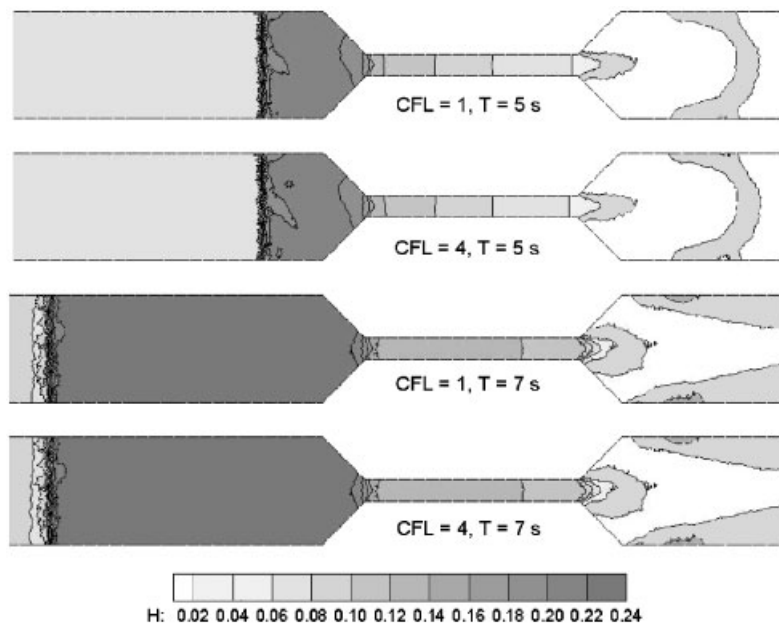


Figure 25. Water levels, in meters, at times $t = 5$, and 7 s, for values of $CFL = 1$ and 4 .

variable bottom slope. The assumed bed level z is defined by a hypothetical function of X , $X = (x, y)$ in meters as

$$z(X) = -0.15 + 0.075 \sin \phi_1 + 0.05 \sin \phi_2$$

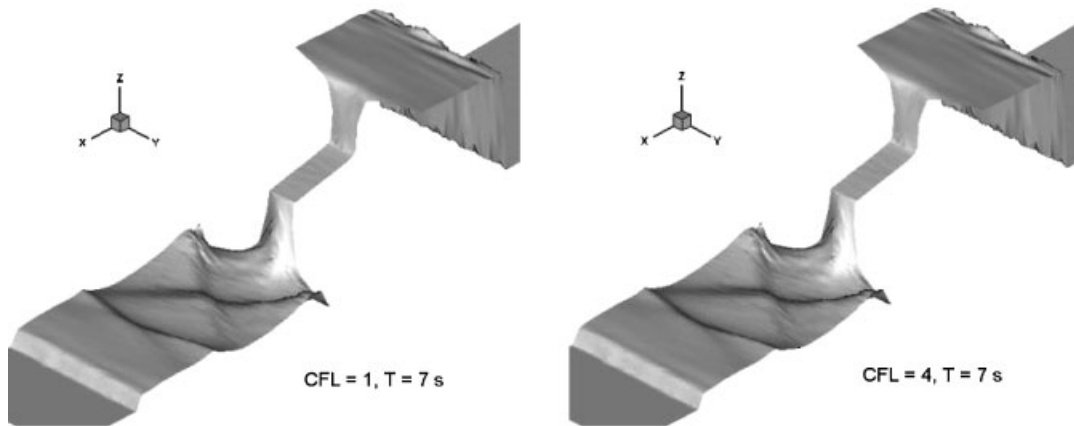


Figure 26. Surface water levels at time 7 s, computed with CFL = 1 and 4.

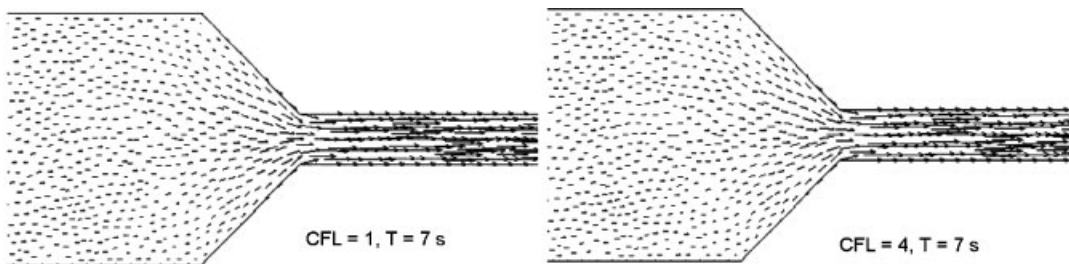


Figure 27. Velocity vector maps at time 7 s, computed with CFL = 1 and 4.

where $X_0 = (0, 0.25)$, $\phi_1 = 4\pi\|X - X_0\|$, $\delta y = (y - 0.25)$ and $\phi_2 = \pi|\delta y|\sin(\phi_1)$. In this case the discontinuity is located closer to the throat of the channel, at $x = 10$ m. Figure 29 shows a detail of the bed in the throat region. The bed level variations exaggerate the 2D character of the flow in the channel.

The existence of a bed slope changes the properties of the flow and the range of the maximum allowed CFL value, but the same tendencies are found. Figure 30 shows the variation of the time step in function of time (left), for target CFL values 1, 2, 3, and 4, enforcing condition (72). The results imposed using condition (73) are equal for the target CFL values 1, 2 and 3, but for values greater than 3 the maximum allowable time step does not increase. The iso-level maps in Figure 31, the 3D images in Figure 32, and the vector velocity maps in Figure 33, show again almost imperceptible differences.

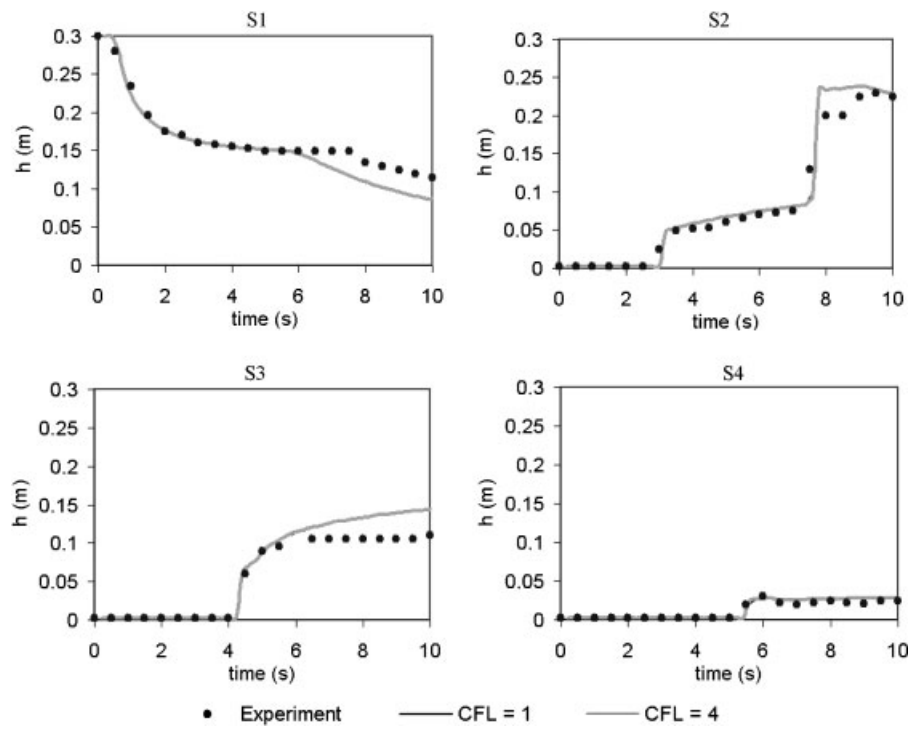


Figure 28. Comparison between experimental data and numerical data for CFL = 1 and 4.

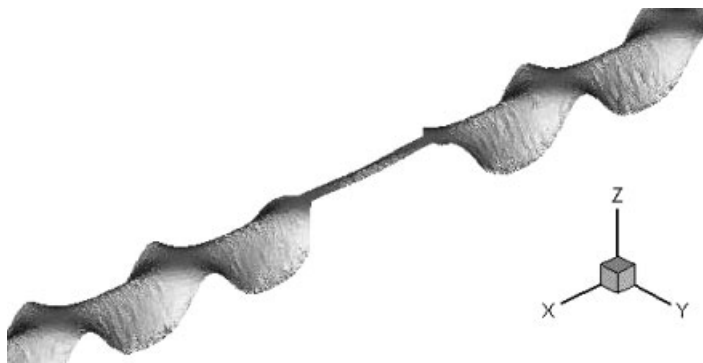


Figure 29. Channel bottom.

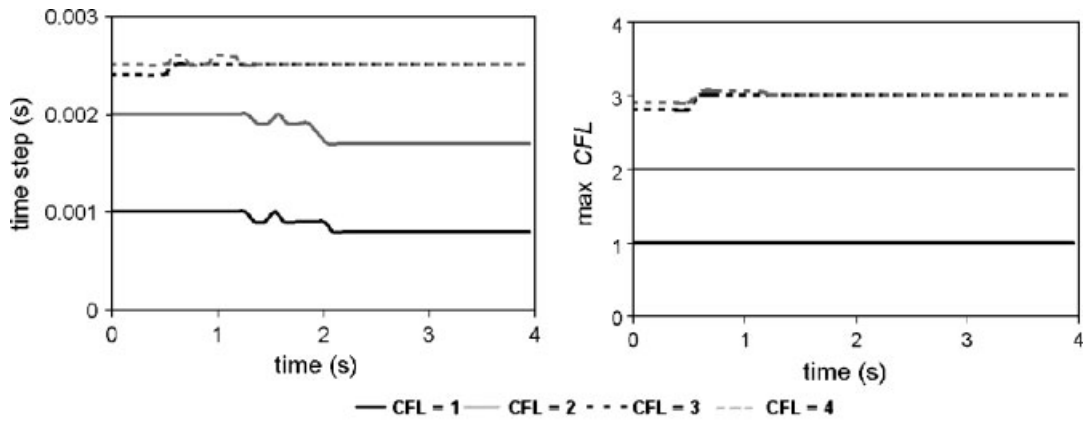


Figure 30. Evolution of the time step (left) and maximum allowable CFL (right) in function of time using (72).

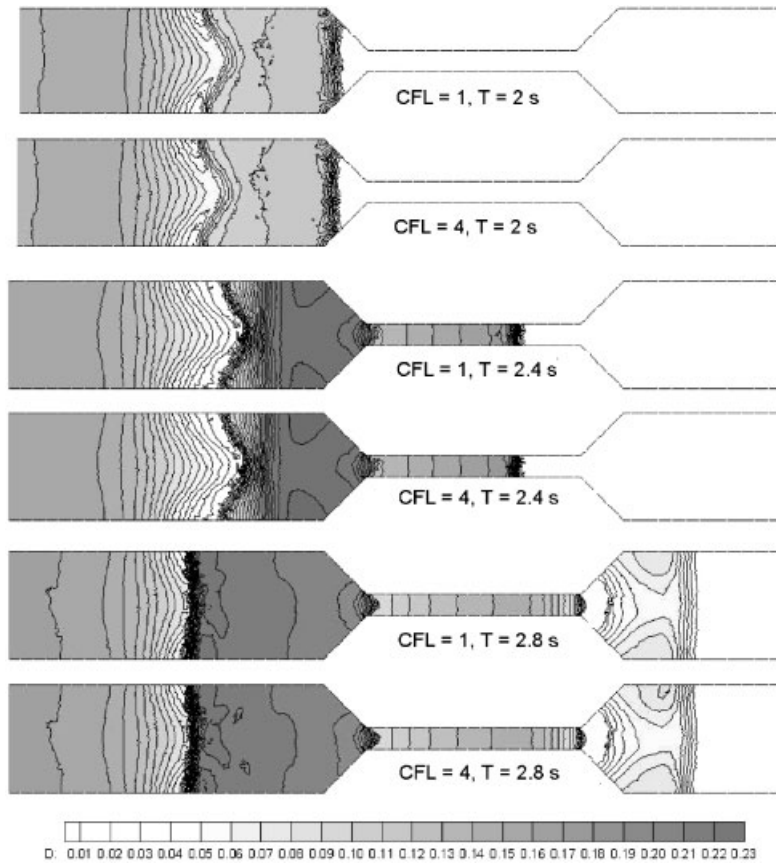


Figure 31. Water surface levels at times $t=2, 2.4$ and 2.8 s, for values of $CFL=1$ and 4 .

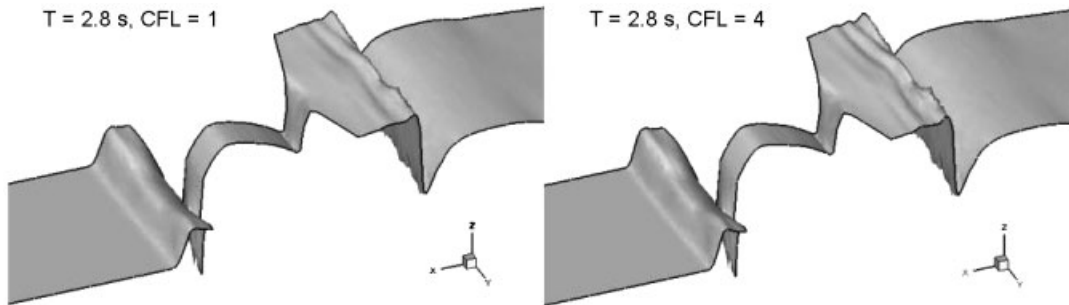


Figure 32. Surface water levels at time 2.8 s, computed with CFL = 1 and 4.

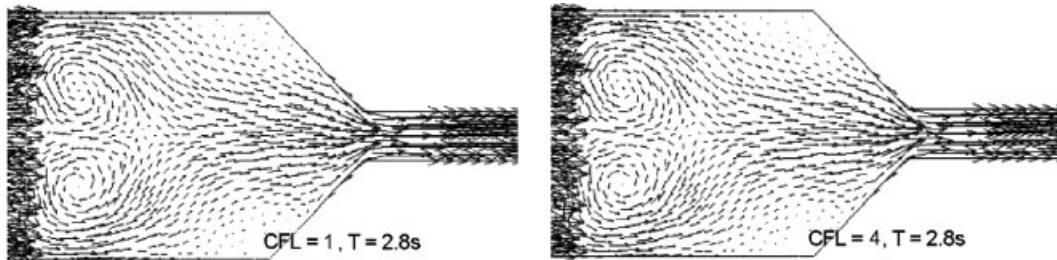


Figure 33. Velocity vector maps at time 2.8 s, computed with CFL = 1 and 4.

3. CONCLUSIONS

An extension of the first order explicit upwind scheme to CFL values greater than 1 has been presented. It has been developed for one and two-dimensional models and applied to linear and non-linear problems discretized on irregular grids. The basic idea linked to the concept of the CFL restrictions for numerical stability in explicit schemes has been revisited. The CFL value is directly related to the extent of the stencil of grid cells involved in one time step. An existing idea of formulating upwind schemes as wave propagations arising from discontinuities in the variables at cell interfaces has been generalized. From the analysis performed, a first conclusion is that it is very important to properly define in each case the reference time step used to state the dimensionless CFL number and that it must always be estimated on the basis of the mentioned stencil. This is crucial when moving from the basic linear case onto regular grids to any other case.

The second important conclusion of our work is that, as the cell interface waves or signals are distributed to more and more neighbour cells as the CFL increases, the numerical solution is extremely grid dependent. This dependence leads to great errors when using structured grids in 2D applications. Hence, the present tool must be used on anisotropic unstructured grids. This can be considered a weakness of the method since a suitable grid generator must be available.

The resulting technique has been applied to transient linear examples in 1D and 2D as well as to steady and unsteady non-linear shallow water test cases with and without source terms.

The results obtained indicate that this method is a useful and robust tool easy to implement as an extension of a previous explicit model of the upwind type in order to relax the stability restrictions over the time step size. The quality of the results is also useful to demonstrate that the linear superposition of waves in which the method is based is a good approximation and does not lead to errors in non-linear cases. The numerical results show that the technique is able to reproduce non-linear flow discontinuities with $CFL > 1$ as efficiently as it does with $CFL < 1$.

Some numerical difficulties have been encountered at solid wall boundaries when using high CFL values. This can be considered a second disadvantage of the method but the way to overcome it is envisaged. The results that can be obtained have been illustrated by means of the application to the scalar advection equation and the shallow water systems with source terms both in one and two-dimensions. Actual work in preparation deals with an exhaustive identification of the behaviour of the numerical solutions provided by this method and those obtained with second order schemes as well as the influence of grid refinement.

REFERENCES

1. Roe PL. Approximate Riemann solvers, parameter vectors and difference schemes. *Journal of Computational Physics* 1981; **43**:357–372.
2. Brufau P, Vázquez-Cedón ME, García-Navarro P. A numerical model for the flooding and drying of irregular domains. *International Journal for Numerical Methods in Fluids* 2002; **39**:247–275.
3. 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* 2001; **37**:209–248.
4. Hubbard ME, García-Navarro P. Flux difference splitting and the balancing of source terms and flux gradients. *Journal of Computational Physics* 2000; **165**:89–125.
5. Toro EF. *Shock-Capturing Methods for Free-Surface Shallow Flows*. Wiley: New York, 2000.
6. Courant R, Isaacson E, Rees M. On the solution of nonlinear hyperbolic differential equations by finite differences. *Communications on Pure and Applied Mathematics* 1952; **5**:243–255.
7. Leveque RJ. Large time step shock-capturing techniques for scalar conservation laws. *Numerical Analysis Project, Manuscript NA-81-13*, Stanford University, Stanford, CA, 1981.
8. Leveque RJ. A large time step generalization of Godunov's method for systems of conservation laws. *SIAM Journal on Numerical Analysis* 1985; **33**:1051–1073.
9. Guinot V. The time line interpolation method for large time step Godunov type schemes. *Journal of Computational Physics* 2002; **130**:1–24.
10. Guinot V. An unconditionally stable, explicit Godunov scheme for systems of conservation laws. *International Journal for Numerical Methods in Fluids* 2002; **38**:567–588.
11. Rosinger EE. Optimal accuracy of unconditionally stable explicit numerical methods for nonlinear evolution PDEs. *Applied Mathematics Letters* 1994; **7**(5):97–100.
12. Rosinger EE, Van Niekerk FD. Application of a general accurate unconditionally stable explicit method to linear and nonlinear Schrödinger equations. *Communications on Applied Nonlinear Analysis* 1994; **1**(4):89–108.
13. Harten A, Hyman P. Self adjusting grid methods for one-dimensional hyperbolic conservation laws. *Journal of Computational Physics* 1982; **50**:235–269.
14. Roe PL. Upwind differencing schemes for hyperbolic conservation laws with source terms. *Proceedings of the 1st International Congress on Hyperbolic Problems*. Springer: Heidelberg, 1986; 41–51.
15. García-Navarro P. Estudio de la aplicabilidad de un esquema explícito descentrado a problemas 1D no estacionarios. *XV C.E.D.Y.A.*, Vigo, Spain, 1997; 442–447.
16. Leveque RJ. *Numerical Methods for Conservation Laws*. Birkhäuser: Basel, 1990.
17. MacDonald I. Analysis and computation of steady open channel flow. *Ph.D. Thesis*, University of Reading, Reading, MA, 1996.
18. George PL, Borouchaki H. *Delaunay Triangulation and Meshing*. Hermes Ed. 1998.
19. Méndez L, Pena P, Brufau, Puertas J. An experimental approach to the dam break problem. Real data to check numerical models. *XXIX IAHR Congress*, Beijing, China, 2001.

20. Soares-Frazao S, Alcrudo F. Conclusions from the 3rd Meeting of the IAHR Working Group on Dam-Break Modelling. CADAM Ed. European Commission, 1999.
21. Billet SJ, Toro EF. On WAF-type schemes for multidimensional hyperbolic conservation laws. *Journal of Computational Physics* 1997; **130**:1–24.