

Implicit schemes with large time step for non-linear equations: application to river flow hydraulics

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

Fluid Mechanics, CPS, University of Zaragoza, 50018 Zaragoza, Spain

SUMMARY

In this work, first-order upwind implicit schemes are considered. The traditional tridiagonal scheme is rewritten as a sum of two bidiagonal schemes in order to produce a simpler method better suited for unsteady transcritical flows. On the other hand, the origin of the instabilities associated to the use of upwind implicit methods for shock propagations is identified and a new stability condition for non-linear problems is proposed. This modification produces a robust, simple and accurate upwind semi-explicit scheme suitable for discontinuous flows with high Courant–Friedrichs–Lewy (CFL) numbers.

The discretization at the boundaries is based on the condition of global mass conservation thus enabling a fully conservative solution for all kind of boundary conditions.

The performance of the proposed technique will be shown in the solution of the inviscid Burgers' equation, in an ideal dambreak test case, in some steady open channel flow test cases with analytical solution and in a realistic flood routing problem, where stable and accurate solutions will be presented using CFL values up to 100. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: implicit schemes; upwind discretization; shallow water; unsteady flow with shocks; non-linear equations; high CFL number

1. INTRODUCTION

Implicit schemes are well known for the property of allowing numerical stability in the resolution of partial differential equations in presence of time steps not restricted by the Courant–Friedrichs–Lewy (CFL) condition. Therefore, they have traditionally been the most attractive methods in CFD for steady or gradually unsteady flows. Computational hydraulics has been a field of frequent search for accurate and robust implicit schemes [1, 2].

Some research has been recently oriented to the development of new implicit techniques able to deal with transcritical flows in order to overcome deficiencies found in previous implicit methods. Among them, for instance, the tridiagonal scheme of Beam–Warming [3]

* Correspondence to: P. García-Navarro, Fluid Mechanics, CPS, University of Zaragoza, 50018 Zaragoza, Spain.

† E-mail: pigar@posta.unizar.es

‡ E-mail: jburguete@able.es

and the two-step bidiagonal implicit MacCormack scheme [4], extended by Casier *et al.* [5] and applied by Meselhe *et al.* [6] to channel and river flows. These methods, although able to capture transcritical transitions and discontinuities, are not free from numerical oscillations and require the addition of artificial viscosity in order to attenuate them. The implicit first-order upwind scheme proposed by Yee [7] was the first non-oscillatory shock capturing implicit scheme.

The results from the mentioned schemes are, however, characterized by a limitation in the allowable time step size in cases of unsteady and discontinuous flow. It is usually found that CFL values bigger than 3 lead to instabilities in the resolution of moving front waves. This is due to a poor treatment of the linearization of the implicit flux terms, which are often locally evaluated.

In this work, new upwind implicit schemes are presented driven by the interest in accurate and efficient river flow modelling. In the line of previous upwind implicit schemes, a new version of the tridiagonal implicit upwind scheme is proposed expressing it as a sum of two bidiagonal schemes. This modification makes it simpler and slightly better adapted to the solution of unsteady transcritical problems.

The application of high CFL numbers when using a conservative implicit scheme for the solution of a non-linear problem with discontinuities originates a special kind of instability at the front position. This is analysed in Appendix B and a new condition is defined so that the numerical schemes do not amplify discontinuities in time (no discontinuity amplification condition, NDA). The NDA condition is applied to a new semi-explicit bidiagonal scheme in order to remove the above-mentioned stability restriction and to produce a simple, robust and accurate upwind scheme suitable for discontinuous flows in complex topography and able to cope with high CFL numbers. The performance of the proposed technique will be shown in several examples going from test cases using the non-linear Burgers' equation to a river flood case of practical interest.

2. THE EQUATIONS

We are interested in solving as efficiently as possible 1D hyperbolic systems with source terms. In a general conservative form

$$\frac{\partial \mathbf{u}(x, t)}{\partial t} + \frac{d\mathbf{F}(x, \mathbf{u})}{dx} = \mathbf{H}(x, \mathbf{u}) \quad (1)$$

where \mathbf{u} is the vector of conserved variables, \mathbf{F} the vector of fluxes and \mathbf{H} the vector of source terms. Our interest is led by the numerical modelling of one-dimensional shallow water flows of practical application in Hydraulics such as river flows. In that case

$$\mathbf{u} = \begin{pmatrix} A \\ Q \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} Q \\ \frac{Q^2}{A} + gI_1 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 0 \\ g[I_2 + A(S_0 - S_f)] \end{pmatrix}$$

where Q is the discharge, A is the wetted cross-section, g is the acceleration of gravity and S_0 is the bed slope. The rest of the terms account for pressure forces and for friction forces, with S_f associated to wall friction and represented by the empirical Manning law [8].

From the equations in conservative form (1), it is possible to pass to an associated non-conservative form

$$\frac{\partial \mathbf{u}(x, t)}{\partial t} + \mathbf{J}(x, \mathbf{u}) \frac{\partial \mathbf{u}(x, t)}{\partial x} = \mathbf{H}'(x, \mathbf{u}) \tag{2}$$

where $\mathbf{J} = \partial \mathbf{F} / \partial \mathbf{u}$ is the Jacobian matrix of the original system and $\mathbf{H}' = \mathbf{H} - \partial \mathbf{F} / \partial x$. In shallow water equations

$$\mathbf{H}' = \begin{pmatrix} 0 \\ gA (S_0 - S_f - \frac{dh}{dx} + \frac{1}{B} \frac{dA}{dx}) \end{pmatrix}, \quad \mathbf{J} = \begin{pmatrix} 0 & 1 \\ c^2 - u^2 & 2u \end{pmatrix}$$

with h the depth, $u = Q/A$ the fluid velocity, $c = \sqrt{gA/B}$ the velocity of the infinitesimal surface waves and B the top width.

It is now convenient to develop the characteristic form of the equations given the importance it has for the correct formulation of upwind schemes and boundary conditions. This form is obtained from a diagonalization of the Jacobian in (2). Calling \mathbf{P} and \mathbf{P}^{-1} the matrices that make \mathbf{J} diagonal, and $\mathbf{\Lambda}$ the resulting diagonal matrix

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

The matrix $\mathbf{\Lambda}$ is formed by the eigenvalues of \mathbf{J} , and \mathbf{P} is constructed with its eigenvectors. Let \mathbf{w} be the set of variables (characteristic variables) that verify

$$d\mathbf{u} = \mathbf{P} d\mathbf{w}, \quad d\mathbf{w} = \mathbf{P}^{-1} d\mathbf{u}$$

Then,

$$\frac{\partial \mathbf{w}(x, t)}{\partial t} + \mathbf{\Lambda}(x, \mathbf{w}) \frac{\partial \mathbf{w}(x, t)}{\partial x} = \mathbf{P}^{-1}(x, \mathbf{w}) \mathbf{H}'(x, \mathbf{w}) \tag{3}$$

In the one-dimensional shallow water model, the above matrices are

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ u + c & u - c \end{pmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} u + c & 0 \\ 0 & u - c \end{pmatrix}$$

3. CONSERVATIVE SCHEMES

The most common definition of a conservative scheme follows the structure [9]

$$\Delta \mathbf{u}_i^n = \Delta t \left[\mathbf{H}_i^* - \frac{1}{\delta x} (\mathbf{F}_{i+1/2}^* - \mathbf{F}_{i-1/2}^*) \right] \tag{4}$$

where \mathbf{H}^* is the numerical source at cell i and $\mathbf{F}_{i+1/2}^*$ is the numerical flux at cell interface $i + 1/2$. They represent a suitable approximation of the true source and flux terms in the equation. Δ will be used for time increment $\Delta f_i^n = f_i^{n+1} - f_i^n$, and δ represents spatial increment $\delta f_{i+1/2} = f_{i+1}^n - f_i^n$. Schemes so defined will be conservative since they cancel the contributions of the flux at the grid interfaces, being the global variation of the conserved variable due only to the source terms and to the fluxes at the boundaries.

A numerical flux \mathbf{F}_i^T can also be defined at the grid nodes [10]. The difference in this flux between two nodes can be decomposed into parts affecting the nodes on the left and right. Schemes so built follow

$$\begin{aligned}\delta\mathbf{F}_{i+1/2}^T &= \mathbf{F}_{i+1}^T - \mathbf{F}_i^T = \delta\mathbf{F}_{i+1/2}^R + \delta\mathbf{F}_{i+1/2}^L \\ \Delta\mathbf{u}_i^n &= \Delta t \left[\mathbf{H}_i^* - \frac{1}{\delta x} (\delta\mathbf{F}_{i+1/2}^R + \delta\mathbf{F}_{i-1/2}^L) \right]\end{aligned}\quad (5)$$

This also leads to conservative schemes since this form can be shown to be equivalent to (4) [10].

In addition, and following evidences from previous works [10–13], we consider a non-pointwise contribution of the source terms

$$\mathbf{H}_{i+1/2}^T = \mathbf{H}_{i+1/2}^R + \mathbf{H}_{i+1/2}^L$$

so that the following formulation for the conservative scheme is proposed:

$$\Delta\mathbf{u}_i^n = \Delta t \left[\left(\mathbf{H} - \frac{\delta\mathbf{F}}{\delta x} \right)_{i-1/2}^L + \left(\mathbf{H} - \frac{\delta\mathbf{F}}{\delta x} \right)_{i+1/2}^R \right]\quad (6)$$

Conservative schemes can also be derived from the non-conservative form of the equations (2) [10]. The advantage is that the latter form tends to be simpler to deal with than the conservative form. We need to establish the conditions under which schemes derived in this way are equivalent to the conservative scheme (6) derived from the conservative equations. First of all, the following equality must hold at the discrete level

$$\mathbf{G}_{i+1/2} \equiv \left(\mathbf{H} - \frac{\delta\mathbf{F}}{\delta x} \right)_{i+1/2} = \left(\mathbf{H}' - \mathbf{J} \frac{\delta\mathbf{u}}{\delta x} \right)_{i+1/2}\quad (7)$$

Note that this equality requires a non-pointwise treatment of source terms and is an extension of the Roe's average [14]. In shallow water this average is [13]

$$\tilde{u}_{i+1/2} = \frac{Q_{i+1}/\sqrt{A_{i+1}} + Q_i/\sqrt{A_i}}{\sqrt{A_{i+1}} + \sqrt{A_i}}, \quad \tilde{c}_{i+1/2} = \sqrt{g \frac{A_{i+1/2}}{B_{i+1/2}}}$$

From (7) it follows that two equivalent forms of building conservative schemes with non-pointwise source terms are possible [10]. Defining \mathbf{G} as

$$\mathbf{G}_{i+1/2} \equiv \left(\mathbf{H} - \frac{\delta\mathbf{F}}{\delta x} \right)_{i+1/2}\quad (8)$$

form (6) is achieved, whereas defining this term like

$$\mathbf{G}_{i+1/2} \equiv \left(\mathbf{H}' - \mathbf{J} \frac{\delta\mathbf{u}}{\delta x} \right)_{i+1/2}\quad (9)$$

with restriction (7), the non-conservative form of the conservative scheme is derived. In any case

$$\Delta\mathbf{u}_i^n = \Delta t (\mathbf{G}_{i-1/2}^L + \mathbf{G}_{i+1/2}^R)\quad (10)$$

where the decomposition in left and right parts is to be defined according to every particular numerical scheme.

Conservative schemes based in the characteristic form of the equations are the basis for the wave decomposition of upwind schemes. From (3) it is possible to rewrite

$$\frac{\partial \mathbf{w}}{\partial t} = \mathbf{P}^{-1} \left(\mathbf{H}' - \mathbf{J} \frac{\partial \mathbf{u}}{\partial x} \right) = \mathbf{P}^{-1} \mathbf{G}$$

Then, a discrete wave decomposition into left and right moving contributions can be done [10]

$$(\mathbf{P}^{-1} \mathbf{G})_{i+1/2} = (\mathbf{\Omega}^L \mathbf{P}^{-1} \mathbf{G})_{i+1/2} + (\mathbf{\Omega}^R \mathbf{P}^{-1} \mathbf{G})_{i+1/2}$$

being $\mathbf{\Omega}^L$ and $\mathbf{\Omega}^R$ diagonal matrices to be defined in every particular numerical scheme. In order to ensure the conservative character of the scheme, they have to obey

$$(\mathbf{\Omega}^L + \mathbf{\Omega}^R)_{i+1/2} = \mathbf{I} \tag{11}$$

Multiplication by \mathbf{P} gives the final form for the discretization in terms of the conserved quantities

$$\Delta \mathbf{u}_i^n = \Delta t [(\mathbf{P} \mathbf{\Omega}^L \mathbf{P}^{-1} \mathbf{G})_{i-1/2} + (\mathbf{P} \mathbf{\Omega}^R \mathbf{P}^{-1} \mathbf{G})_{i+1/2}] \tag{12}$$

Note that this discretization requires again a non-pointwise formulation of the source terms, being equally possible (8) or (9) for \mathbf{G} . If a pointwise treatment of the source terms is desired, the following can be made:

$$\Delta \mathbf{u}_i^n = \Delta t \left[\mathbf{H}_i^* - \left(\mathbf{P} \mathbf{\Omega}^L \mathbf{P}^{-1} \frac{\delta \mathbf{F}}{\delta x} \right)_{i-1/2} - \left(\mathbf{P} \mathbf{\Omega}^R \mathbf{P}^{-1} \frac{\delta \mathbf{F}}{\delta x} \right)_{i+1/2} \right] \tag{13}$$

Both options can be inserted into the general form of the conservative schemes (10) and (5) by definition of

$$\mathbf{G}^{L,R} = \mathbf{P} \mathbf{\Omega}^{L,R} \mathbf{P}^{-1} \mathbf{G}, \quad \delta \mathbf{F}^{L,R} = \mathbf{P} \mathbf{\Omega}^{L,R} \mathbf{P}^{-1} \delta \mathbf{F} \tag{14}$$

4. FIRST-ORDER IMPLICIT UPWIND SCHEMES

4.1. Tridiagonal scheme

Upwind schemes are based on the idea of approximating the spatial derivatives by non-centred differences biased in the sense of propagation of information in the physical problem. In order to construct a first-order scheme, suitable for left and right moving propagation velocities, the following wave decomposition can be made:

$$\mathbf{F}_i^T = \mathbf{F}_i^{n+\theta}, \quad \delta \mathbf{F}_{i+1/2}^L = (\delta \mathbf{F}^+)_{i+1/2}^{n+\theta}, \quad \delta \mathbf{F}_{i+1/2}^R = (\delta \mathbf{F}^-)_{i+1/2}^{n+\theta}$$

leading to

$$\Delta \mathbf{u}_i^n = \Delta t [(\mathbf{G}^+)_{i-1/2}^{n+\theta} + (\mathbf{G}^-)_{i+1/2}^{n+\theta}] \tag{15}$$

where \mathbf{G}^- , $\delta\mathbf{F}^-$ and \mathbf{J}^- are associated to negative velocities, and \mathbf{G}^+ , $\delta\mathbf{F}^+$ and \mathbf{J}^+ to positive velocities and the notation $f^{n+\theta} = \theta f^{n+1} + (1-\theta)f^n$ is used. From (12), the following wave decomposition is assumed for this scheme in order to select the appropriate influence region in every case:

$$\begin{aligned}\mathbf{\Omega}^L &= \mathbf{\Omega}^+ = \frac{1}{2}[\mathbf{I} + \mathbf{\Lambda}^{-1}|\mathbf{\Lambda}|], & \mathbf{\Omega}^R &= \mathbf{\Omega}^- = \frac{1}{2}[\mathbf{I} - \mathbf{\Lambda}^{-1}|\mathbf{\Lambda}|] \\ \mathbf{J}^\pm &= \mathbf{P}\mathbf{\Omega}^\pm\mathbf{P}^{-1}\mathbf{J}, & \delta\mathbf{F}^\pm &= \mathbf{P}\mathbf{\Omega}^\pm\mathbf{P}^{-1}\delta\mathbf{F}, & \mathbf{G}^\pm &= \mathbf{P}\mathbf{\Omega}^\pm\mathbf{P}^{-1}\mathbf{G}\end{aligned}\quad (16)$$

A linear analysis of the homogeneous equations shows that the stability condition is

$$(1 - 2\theta)\text{CFL} \leq 1 \quad (17)$$

with $\text{CFL} = \max |a^k| \Delta t / \delta x$ the CFL number and a^k the eigenvalues of the Jacobian, being unconditionally stable if $\theta \geq \frac{1}{2}$. This linear analysis also shows that the scheme is total variation diminishing (TVD) if

$$\text{CFL} \leq \frac{1}{1 - \theta} \quad (18)$$

being this condition more restrictive. The scheme is unconditionally TVD if $\theta = 1$.

With non-linear equations this scheme requires to involve \mathbf{F}^{n+1} and \mathbf{H}^{n+1} , which represents a difficulty. In order to avoid this problem, the following linearization can be made

$$\begin{aligned}\mathbf{F}^{n+1} &\approx \mathbf{F}^n + \frac{\partial \mathbf{F}^n}{\partial t} \Delta t \approx \mathbf{F}^n + \left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial t} \right)^n \Delta t \approx \mathbf{F}^n + \mathbf{J}^n \Delta \mathbf{u}^n \\ \mathbf{H}^{n+1} &\approx \mathbf{H}^n + \frac{\partial \mathbf{H}^n}{\partial t} \Delta t \approx \mathbf{H}^n + \left(\frac{\partial \mathbf{H}}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial t} \right)^n \Delta t \approx \mathbf{H}^n + \mathbf{K}^n \Delta \mathbf{u}^n\end{aligned}$$

with $\mathbf{K} = \partial \mathbf{H} / \partial \mathbf{u}$ the Jacobian of the source term. Using the wave decomposition (16), the conservative scheme in characteristic form (12) can be expressed

$$\begin{aligned}\theta \frac{\Delta t}{\delta x} (\mathbf{J}^- \Delta \mathbf{u})_{i+1}^n + \left[1 - \theta \Delta t \left(\mathbf{K} - \frac{\mathbf{J}^+}{\delta x} + \frac{\mathbf{J}^-}{\delta x} \right) \right]_i^n \Delta \mathbf{u}_i^n - \theta \frac{\Delta t}{\delta x} (\mathbf{J}^+ \Delta \mathbf{u})_{i-1}^n \\ = \Delta t [(\mathbf{G}^+)_{i-1/2}^n - (\mathbf{G}^-)_{i+1/2}^n]\end{aligned}\quad (19)$$

with $\mathbf{J}^\pm = \mathbf{P}\mathbf{\Omega}^\pm\mathbf{P}^{-1}\mathbf{J}$. It is worth noting that it is simpler to make the upwind treatment of the source terms in the explicit operator keeping a pointwise discretization for \mathbf{K} in the linearized implicit operator.

The numerical scheme, as described in (15) or in the linearized version (19), is unable to deal with transonic problems of the type subcritical to supercritical flow, requiring in these cases the introduction of an artificial viscosity like the one proposed by Harten–Hyman [15].

In the linearized case, the wave decomposition is

$$\begin{aligned} \mathbf{F}_i^T &= \mathbf{F}_i^n + \theta(\mathbf{J}\Delta\mathbf{u})_i^n \\ \delta\mathbf{F}_{i+1/2}^L &= (\delta\mathbf{F}^+)_i^n + \theta\delta(\mathbf{J}^+\Delta\mathbf{u})_{i+1/2}^n - (v\delta\mathbf{u})_{i+1/2}^n \\ \delta\mathbf{F}_{i+1/2}^R &= (\delta\mathbf{F}^-)_i^n + \theta\delta(\mathbf{J}^-\Delta\mathbf{u})_{i+1/2}^n + (v\delta\mathbf{u})_{i+1/2}^n \end{aligned}$$

where we propose to use the following:

$$v_{i+1/2}^n = \max_k (v^k)_{i+1/2}^n \tag{20}$$

with

$$(v^k)_{i+1/2}^n = \begin{cases} \frac{1}{4}[\delta(a^k)_{i+1/2}^n - 2|a^k|_{i+1/2}^n] & \text{if } (a^k)_i^n < 0 \text{ and } (a^k)_{i+1}^n > 0 \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

Therefore, the quantities \mathbf{G}^\pm will be redefined to cope with the mentioned transitions as

$$(\mathbf{G}^\pm)_{i+1/2}^n = \left(\mathbf{P}\Omega^\pm\mathbf{P}^{-1}\mathbf{G} \pm v \frac{\delta\mathbf{u}}{\delta x} \right)_{i+1/2}^n \tag{22}$$

This linearized implicit scheme was developed by Yee [7] in the context of gas dynamics and successfully used for steady problems. The performance for steady problems is highly sensitive to the CFL used during the transient phase. When shocks are present in the solution of a non-linear equation and a high time step is used, the propagation velocities (eigenvalues in the system case) can change strongly between the two time levels. The implicit linearized methods are not unconditionally stable in these cases because the linearization involves the evaluation of the Jacobian at time level n . In the references the test cases presented confirm this behaviour.

4.2. Bidiagonal scheme

The scheme built as in (19) leads, in general, to a linear block tridiagonal system of equations. In order to work with a simpler block bidiagonal system and to avoid spurious errors in the transcritical points, the original scheme (19) can be altered. We propose the following modification by splitting the increments due to positive and negative propagations

$$\begin{aligned} \mathbf{F}_i^T &= \mathbf{F}_i^n + \theta(\mathbf{J}^+\Delta\mathbf{u}^+ + \mathbf{J}^-\Delta\mathbf{u}^-)_i^n \\ \delta\mathbf{F}_{i+1/2}^L &= (\delta\mathbf{F}^+)_i^n + \theta\delta(\mathbf{J}^+\Delta\mathbf{u}^+)_{i+1/2}^n \\ \delta\mathbf{F}_{i+1/2}^R &= (\delta\mathbf{F}^-)_i^n + \theta\delta(\mathbf{J}^-\Delta\mathbf{u}^-)_{i+1/2}^n \end{aligned}$$

and the original scheme is transformed in a two-step block bidiagonal system:

$$\begin{aligned} \left[1 - \theta \frac{\Delta t}{\delta x} (\mathbf{J}^-)\right]_i^n (\Delta \mathbf{u}^-)_i^n + \theta \frac{\Delta t}{\delta x} (\mathbf{J}^- \Delta \mathbf{u}^-)_{i+1}^n &= \Delta t (\mathbf{G}^-)_{i+1/2}^n \\ \left[1 + \theta \frac{\Delta t}{\delta x} (\mathbf{J}^+)\right]_i^n (\Delta \mathbf{u}^+)_{i-1}^n - \theta \frac{\Delta t}{\delta x} (\mathbf{J}^+ \Delta \mathbf{u}^+)_{i-1}^n &= \Delta t (\mathbf{G}^+)_{i-1/2}^n \\ (1 - \theta \Delta t \mathbf{K}_i^n) \Delta \mathbf{u}_i^n &= (\Delta \mathbf{u}^+ + \Delta \mathbf{u}^-)_i^n \end{aligned} \quad (23)$$

When dealing with scalar equations with no source terms, this discretization is always identical to the classical tridiagonal scheme except for transcritical points. Hence, the stability (17) and TVD (18) conditions, deduced for linear cases, are equally valid for this scheme.

Despite the improvement achieved by the splitting philosophy in transcritical flow points, this scheme is still not able to cope with high CFL numbers in presence of moving shocks.

5. SEMI-EXPLICIT NDA FIRST-ORDER UPWIND SCHEME

In a different approach, scheme (23) can be improved by means of a locally defined spatial weighting parameter. A wave decomposition like

$$\begin{aligned} \mathbf{F}_i^T &= \mathbf{F}_i^n + \frac{\delta x}{\Delta t} (\xi^+ \Delta \mathbf{u}^+ - \xi^- \Delta \mathbf{u}^-)_i^n \\ \delta \mathbf{F}_{i+1/2}^L &= (\delta \mathbf{F}^+)^n_{i+1/2} + \frac{\delta x}{\Delta t} \delta (\xi^+ \Delta \mathbf{u}^+)_{i+1/2}^n \\ \delta \mathbf{F}_{i+1/2}^R &= (\delta \mathbf{F}^-)^n_{i+1/2} + \frac{\delta x}{\Delta t} \delta (\xi^- \Delta \mathbf{u}^-)_{i+1/2}^n \end{aligned}$$

produces the following numerical scheme:

$$\begin{aligned} (1 + \xi^-)_i^n (\Delta \mathbf{u}^-)_i^n - (\xi^-)_{i+1}^n (\Delta \mathbf{u}^-)_{i+1}^n &= \Delta t (\mathbf{G}^-)_{i+1/2}^n \\ (1 + \xi^+)_{i-1}^n (\Delta \mathbf{u}^+)_{i-1}^n - (\xi^+)_{i-1}^n (\Delta \mathbf{u}^+)_{i-1}^n &= \Delta t (\mathbf{G}^+)_{i-1/2}^n \\ (1 - \theta \Delta t \mathbf{K}_i^n) \Delta \mathbf{u}_i^n &= (\Delta \mathbf{u}^+ + \Delta \mathbf{u}^-)_i^n \end{aligned} \quad (24)$$

It is worth noting that this definition of the parameters ξ modifies the wave decomposition of the implicit operator in a manner very similar to the introduction of the artificial viscosity applied in the explicit operator to correct the sub-super transitions. This suggests that the physical meaning of these parameters is analogous and that the ξ 's represent an implicit artificial viscosity.

From a non-linear analysis based on a scalar equation, as described in Appendix B, it is easy to deduce that conditions (B2) and (B3) leading to an NDA scheme can be written as follows for the present scheme:

$$\begin{aligned} \xi_i^\pm &\geq \pm \frac{\Delta t}{\delta x} a_{i \mp \frac{1}{2}}^\pm - 1 \\ \xi_i^\pm &\geq \xi_{i \mp 1}^\pm - 1 \end{aligned}$$

From these conditions, the set of parameters ξ will be calculated as

$$\xi_i^\pm = \max \left[\pm \frac{\Delta t}{\delta x} a_{i \mp \frac{1}{2}}^\pm - 1, \xi_{i \mp 1}^\pm - 1, 0 \right] \tag{25}$$

In those cases of transition from subcritical to supercritical flow, an artificial viscosity needs also to be included and (22) must be applied. It can be easily verified through a detailed non-linear NDA analysis like the one discussed in Appendix B, that the artificial viscosity defined like in (21) also modifies condition (B2) in the following way:

$$\begin{aligned} \xi_i^+ &\geq \frac{\Delta t}{\delta x} a_i^+ - 1 \\ \xi_i^- &\geq -\frac{\Delta t}{\delta x} a_i^- - 1 \end{aligned}$$

Therefore, in these cases the parameters ξ should be defined as

$$\begin{aligned} \xi_i^+ &= \max \left[\frac{\Delta t}{\delta x} a_i^+ - 1, \xi_{i-1}^+ - 1, 0 \right] \\ \xi_i^- &= \max \left[-\frac{\Delta t}{\delta x} a_i^- - 1, \xi_{i+1}^- - 1, 0 \right] \end{aligned} \tag{26}$$

6. BOUNDARY CONDITIONS

The theory of characteristics supplies a rigorous rule for the numerical treatment of the boundary conditions at the upstream and downstream ends. In case of subcritical flow, at every boundary one external or physical boundary condition is required and the numerical scheme must provide another one. On the other hand, if the flow is supercritical, two external boundary conditions need to be imposed at the inlet (upstream) whereas the numerical scheme will provide two numerical boundary conditions at the outlet (downstream) [9, 16].

The method of global conservation proposed here is based on a very important physical principle: the increment of mass in the whole system in a time interval is the result of the entering mass flow minus the leaving mass flow during that period of time. When using a conservative numerical scheme, this physical principle provides a way to reach two objectives, first to allow for a null mass balance error and, second, to supply numerical boundary conditions.

In a first step, the numerical solution \mathbf{u}_i^* is obtained assuming derivatives to be nil at the boundaries. This step involves a numerical (positive or negative) error in the global mass which must be compensated with the global error induced by the physical boundary conditions in order to achieve the global mass conservation in our system [17, 18]. In total, from the mathematical point of view, the external boundary conditions and the requirement of global mass conservation supply enough equations to obtain the solution \mathbf{u}_i^{n+1} leading to conservative solutions up to machine accuracy.

Let us assume a problem discretized by means of N cells:

$$M = \sum_{i=1}^N A_i \delta x \quad (27)$$

so that the total mass increment in one time step ΔM^n is

$$\Delta M^n = \sum_{i=1}^N \Delta A_i^n \delta x \quad (28)$$

Suppose we are using a conservative scheme for the interior points with nodal flux as in (10):

$$\Delta \mathbf{u}_i^* = \Delta t (\mathbf{G}_{i-1/2}^L + \mathbf{G}_{i+1/2}^R) \quad (29)$$

although the method is also valid with nodal source term. The assumption of zero derivatives at the boundaries leads to

$$\mathbf{G}_{N+1/2}^R = \mathbf{G}_{1/2}^L = 0, \quad \mathbf{u}_1^* = \mathbf{u}_1^n + \Delta t \mathbf{G}_{3/2}^R, \quad \mathbf{u}_N^* = \mathbf{u}_N^n + \Delta t \mathbf{G}_{N-1/2}^L \quad (30)$$

The wetted cross-section increments at every node in that time step are

$$\Delta A_i^* = -\frac{\Delta t}{\delta x} (\delta Q_{i+1/2}^R + \delta Q_{i-1/2}^L) \quad (31)$$

with

$$\delta Q_{i+1/2}^R + \delta Q_{i+1/2}^L = Q_{i+1}^T - Q_i^T \quad (32)$$

Using (28) the numerical mass variation produced by the scheme is therefore:

$$\Delta M^* = -\Delta t \sum_{i=1}^N (\delta Q_{i-1/2}^L + \delta Q_{i+1/2}^R) \quad (33)$$

Applying (30) and rearranging indexes, the mass variation is

$$\Delta M^* = \Delta t (Q_1^T - Q_N^T) \quad (34)$$

It comes from the approximation at the boundaries and can be split into numerical mass errors at the inlet ΔM_{in}^* and at the outlet ΔM_{out}^* as follows:

$$\Delta M_{\text{in}}^* = \Delta t Q_1^T, \quad \Delta M_{\text{out}}^* = -\Delta t Q_N^T \quad (35)$$

In order to ensure stability when using implicit schemes, perturbations such as those introduced at the boundaries must be allowed to propagate all over the domain. Using a formulation based on the upwind semi-explicit bidiagonal scheme described in 5, can be written as follows:

$$\begin{aligned} \Delta \mathbf{u}_i^+ &= \left(\frac{\xi}{1+\xi} \right)^{i-1} \Delta \mathbf{u}_1^a \\ \Delta \mathbf{u}_i^- &= \left(\frac{\xi}{1+\xi} \right)^{N-i} \Delta \mathbf{u}_N^a \\ \Delta \mathbf{u}_i^n &= \Delta \mathbf{u}_i^* + \Delta \mathbf{u}_i^a = \Delta \mathbf{u}_i^* + \Delta \mathbf{u}_i^+ + \Delta \mathbf{u}_i^- \end{aligned} \quad (36)$$

being $\Delta \mathbf{u}_i^a$ the propagated modifications so that using (28) and the formula of the sum of the terms in a geometric progression, the mass variations at the inlet and outlet are

$$\begin{aligned} \Delta M_{\text{in}}^a &= \delta x \sum_{i=1}^N \Delta A_i^+ = \eta \Delta A_1^a \delta x \\ \Delta M_{\text{out}}^a &= \delta x \sum_{i=1}^N \Delta A_i^- = \eta \Delta A_N^a \delta x \end{aligned} \tag{37}$$

with

$$\eta = (1 + \xi) \left[1 - \left(\frac{\xi}{1 + \xi} \right)^N \right] \tag{38}$$

where one of the simplest and best options for parameter ξ is

$$\xi = \max(\text{CFL} - 1, 0) \tag{39}$$

When the physical boundary conditions imply a real mass increment at the inlet ΔM_{in}^n or at the outlet ΔM_{out}^n , the application of (37) ensures the global mass conservation in the system:

$$\begin{aligned} \Delta M_{\text{in}}^n &= \Delta M_{\text{in}}^* + \Delta M_{\text{in}}^a = \eta \Delta A_1^a \delta x + \Delta t Q_1^T \\ \Delta M_{\text{out}}^n &= \Delta M_{\text{out}}^* + \Delta M_{\text{out}}^a = \eta \Delta A_N^a \delta x - \Delta t Q_N^T \end{aligned} \tag{40}$$

These two equations together with (36) are the basis of the method which will be clarified next with some examples.

6.1. $Q = Q(t)$ at the boundary

When a discharge hydrograph is known at the inlet point, the physical upstream variation is

$$\Delta M_{\text{in}}^n = \int_{t^n}^{t^{n+1}} Q(t) dt \tag{41}$$

hence ΔQ_1^n and ΔQ_1^a are known. Using (40):

$$\Delta A_1^a = \frac{\Delta M_{\text{in}}^n - \Delta t Q_1^T}{\eta \delta x} \tag{42}$$

where it is also necessary to apply (36). The equivalent at the outlet in the case of a known outflow discharge and therefore ΔQ_N^a and ΔM_{out}^n , is

$$\Delta A_N^a = \frac{\Delta M_{\text{out}}^n + \Delta t Q_N^T}{\eta \delta x} \tag{43}$$

6.2. $Q = Q(h)$ at the boundary

In some hydraulic problems, a dynamic stage-discharge $Q = Q(h)$ has to be enforced as external boundary condition. This gives rise to a system of equations that may become difficult

to solve depending on the form of this function. The approximation

$$Q^{n+1} \approx Q(h^n)$$

simplifies the calculation which is then treated as described in the previous subsection.

6.3. $h = h(t)$ at the boundary

When a water level hydrograph or limnigraph is imposed as external boundary condition, either ΔA_1^n and ΔA_1^a at the entrance or ΔA_N^n and ΔA_N^a at the outlet are known. The physical mass variation can be approximated via a second-order trapezoidal rule as

$$\begin{aligned} \Delta M_{\text{in}}^n &= \frac{Q_1^n + Q_1^{n+1}}{2} \Delta t = \left(Q_1^n + \frac{1}{2} \Delta Q_1^n \right) \Delta t \\ \Delta M_{\text{out}}^n &= -\frac{Q_N^n + Q_N^{n+1}}{2} \Delta t = -\left(Q_N^n + \frac{1}{2} \Delta Q_N^n \right) \Delta t \end{aligned} \quad (44)$$

From (40) we can get

$$\Delta Q_1^n = 2 \left(Q_1^T - Q_1^n + \eta \frac{\delta x}{\Delta t} \Delta A_1^a \right) \quad (45)$$

having enforced (36). In the same form, at the outlet, if the water depth and therefore ΔA_N^a are known, it is possible to estimate

$$\Delta Q_N^n = 2 \left(Q_N^T - Q_N^n - \eta \frac{\delta x}{\Delta t} \Delta A_N^a \right) \quad (46)$$

7. NUMERICAL RESULTS

7.1. Burgers' equation

The inviscid Burgers' equation [19]

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} \right) = 0$$

is used to compare the performance of the various implicit upwind schemes in a simple non-linear and homogeneous scalar hyperbolic case with shock. The analytical solutions to this problem can be found in Reference [9].

7.1.1. Unsteady case. For the purpose of having an unsteady solution, initial conditions for the function u are chosen discontinuous and transcritical as depicted in Figure 1. The results from the simulation at time $t = 20$ on a uniform grid with 400 cells are presented and compared with the exact solution in Figures 2 and 3. Both the tridiagonal and bidiagonal schemes can be seen to be unstable from $\text{CFL} = 10$. On the other hand, the semi-explicit NDA scheme

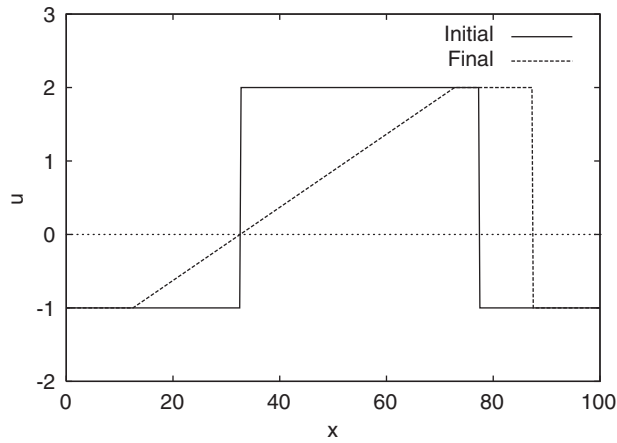


Figure 1. Initial conditions and exact solution for the unsteady Burgers' test case.

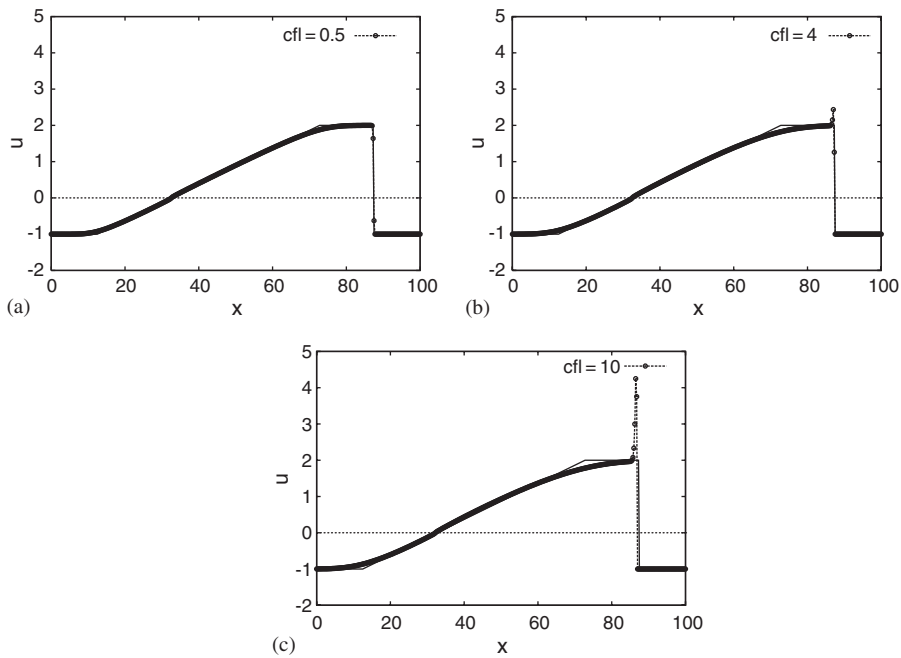


Figure 2. Analytical and numerical solution of the initial square shape using the tridiagonal scheme: (a) CFL = 0.5; (b) CFL = 4; and (c) CFL = 10.

remains unconditionally stable. In Figure 3(f), the numerical solution obtained with the NDA semi-explicit scheme using a CFL = 100 is not smooth; on the contrary, it presents a profile made of jumps that has been produced in only two time steps. In order to verify the sensitivity of the method to the grid refinement as well as the evolution of the shape of the numerical

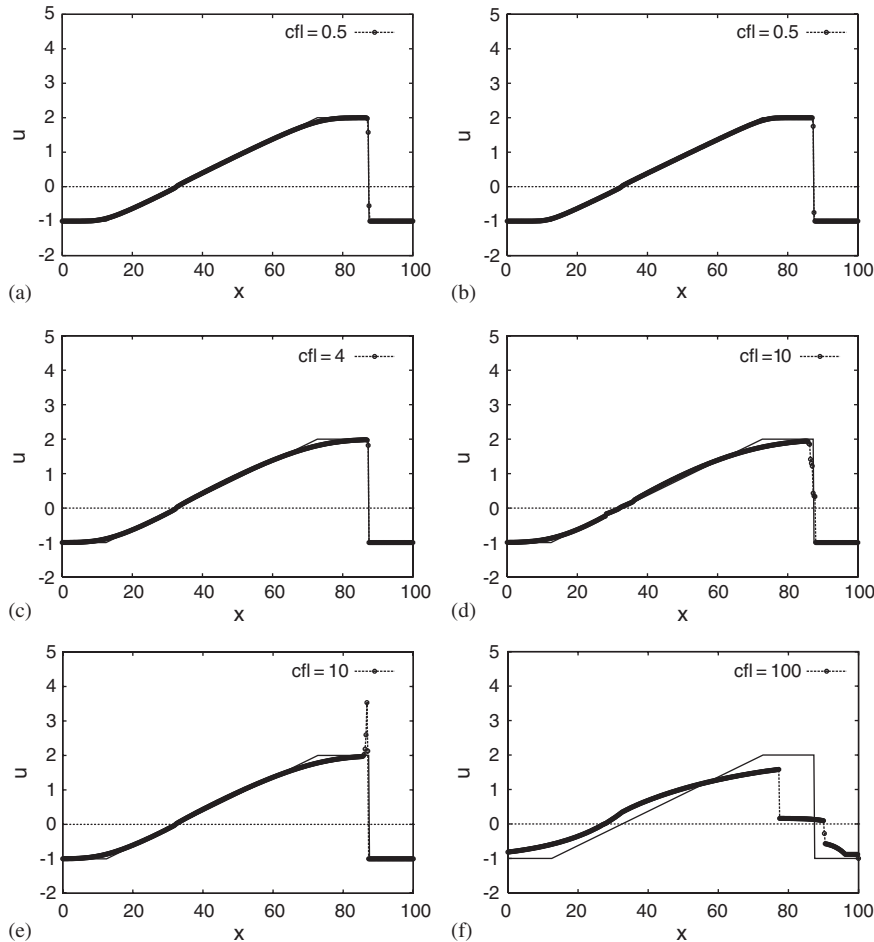


Figure 3. Analytical and numerical solution of the initial square shape using the bidiagonal scheme (a, c, e) and the NDA semi-explicit scheme (b, d, f).

solution, Figure 4 shows the results of the same problem on a finer grid at times $t = 8$ and 20.

7.1.2. Convergence to steady state. The initial conditions shown in Figure 5 have been selected to illustrate that, although the NDA implicit method is able to simulate the unsteady discontinuity solution of a non-linear problem using high CFL number, it may spoil the quality of a discontinuous transcritical steady solution due to an excessive numerical diffusion which grows with the CFL number. This is displayed on Figure 6, where the better performance of the bidiagonal implicit scheme for steady solutions is clear. The reason for this behaviour is that the NDA method promotes the passage of information from non-physical

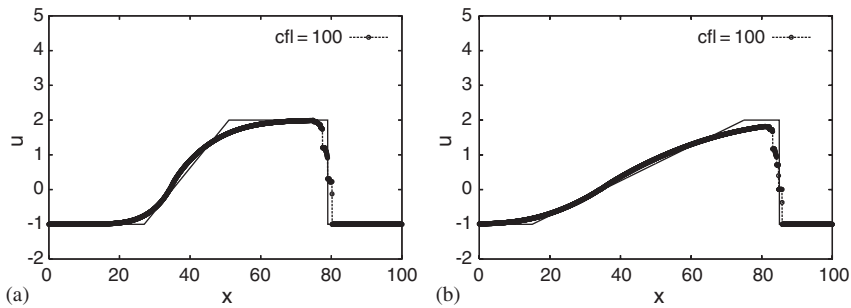


Figure 4. Analytical and numerical solution of the unsteady Burgers' case using the NDA semi-explicit scheme with a finer grid ($\Delta x = 0.04$) at $t = 8$ (a) and $t = 20$ (b).

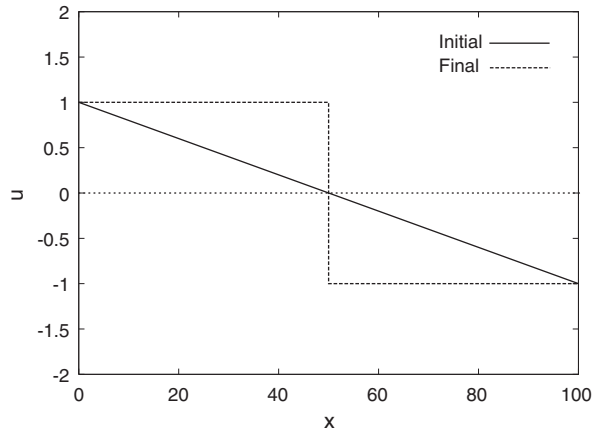


Figure 5. Initial conditions and analytical solution for the steady Burgers' test case.

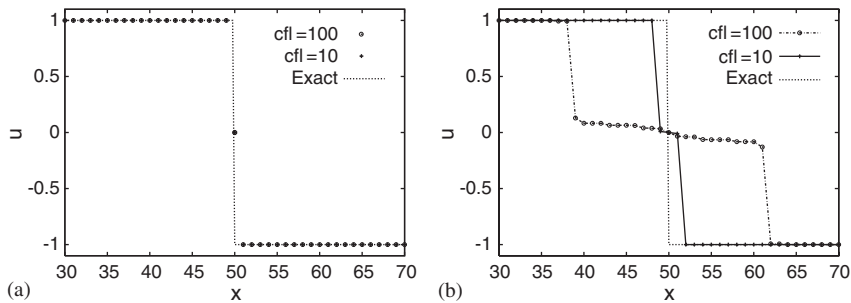


Figure 6. Analytical and numerical solution of the steady Burgers' case using the bidiagonal scheme (a) and the NDA semi-explicit scheme (b).

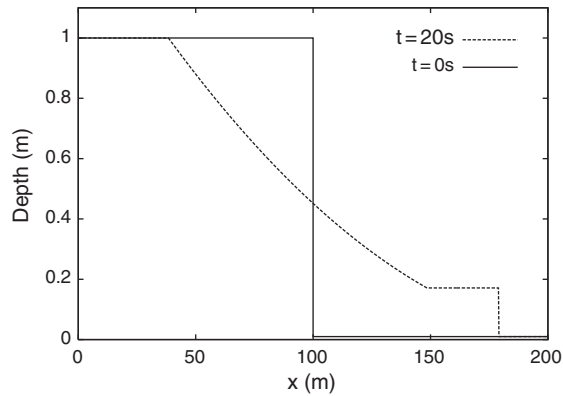


Figure 7. Dambreak test case. Initial conditions and analytical solution.

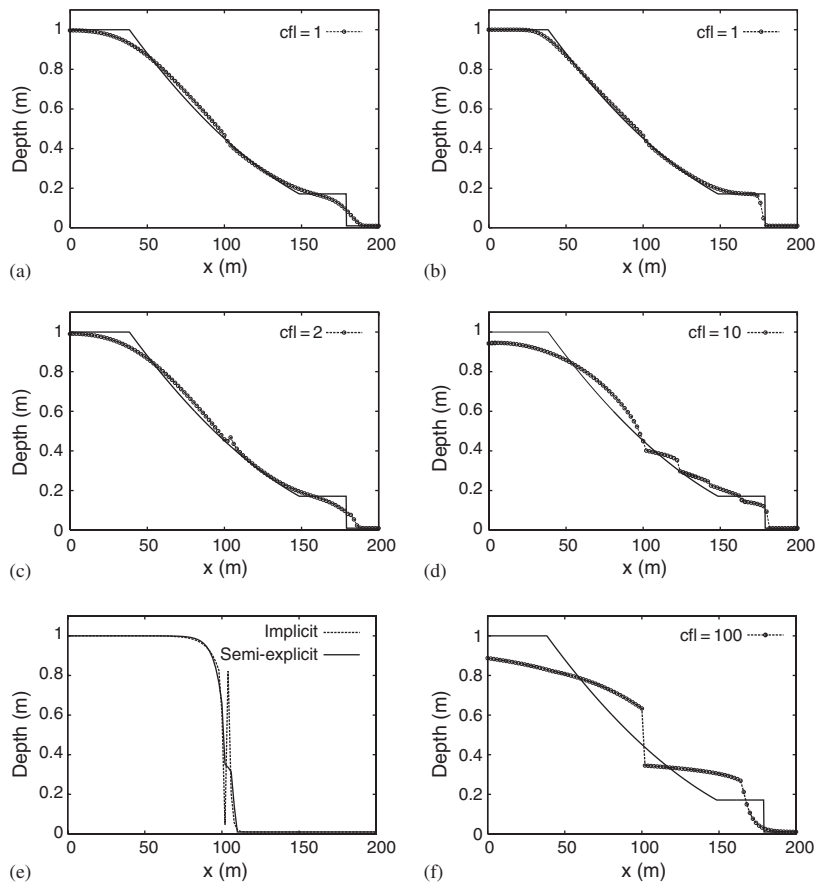


Figure 8. Water depth profiles for the dambreak test case. (a), (c), (e) Bidiagonal implicit. (b), (d), (f) NDA semi-explicit. In (e) both schemes are compared during a single time step using $CFL \approx 6$.

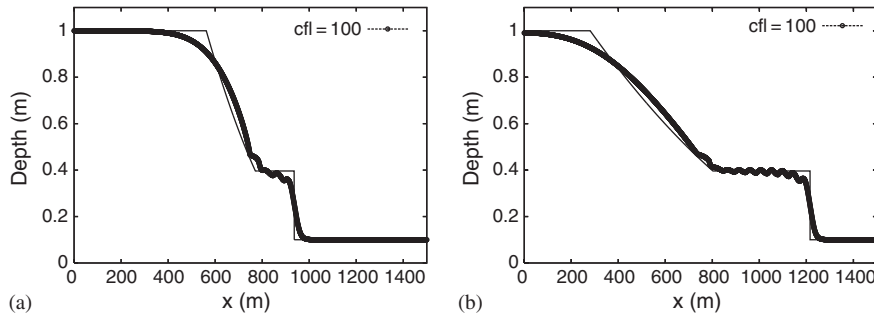


Figure 9. Analytical and numerical solution of the dambreak case using the NDA semi-explicit scheme with CFL = 100 in a longer domain at $t = 60$ s (a) and $t = 150$ s (b).

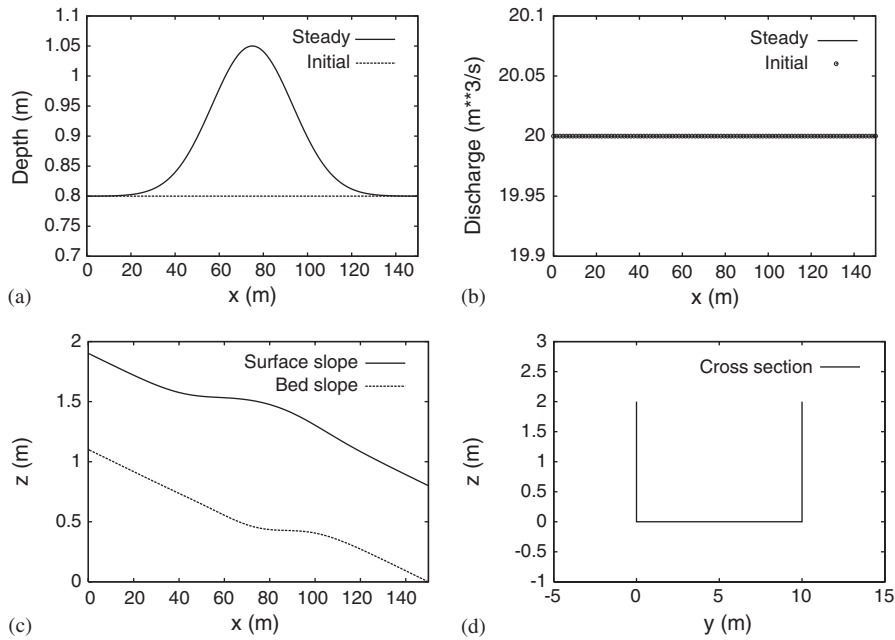


Figure 10. MacDonald 1 test case: (a) Initial and steady water depth profiles; (b) initial and steady discharge profiles; (c) water surface and channel bed profiles; and (d) channel cross-section.

regions of influence across the discontinuity. This is linked to the stabilizing tendency of this method and does not produce oscillatory results.

7.2. Shallow water equations: ideal dam break

The unsteady flow induced by an ideal dam break is the most widely used test case for numerical schemes of the kind considered here. The initial conditions defining this non-linear problem, as well as the exact solution are plotted in Figure 7. Implicit methods are particularly

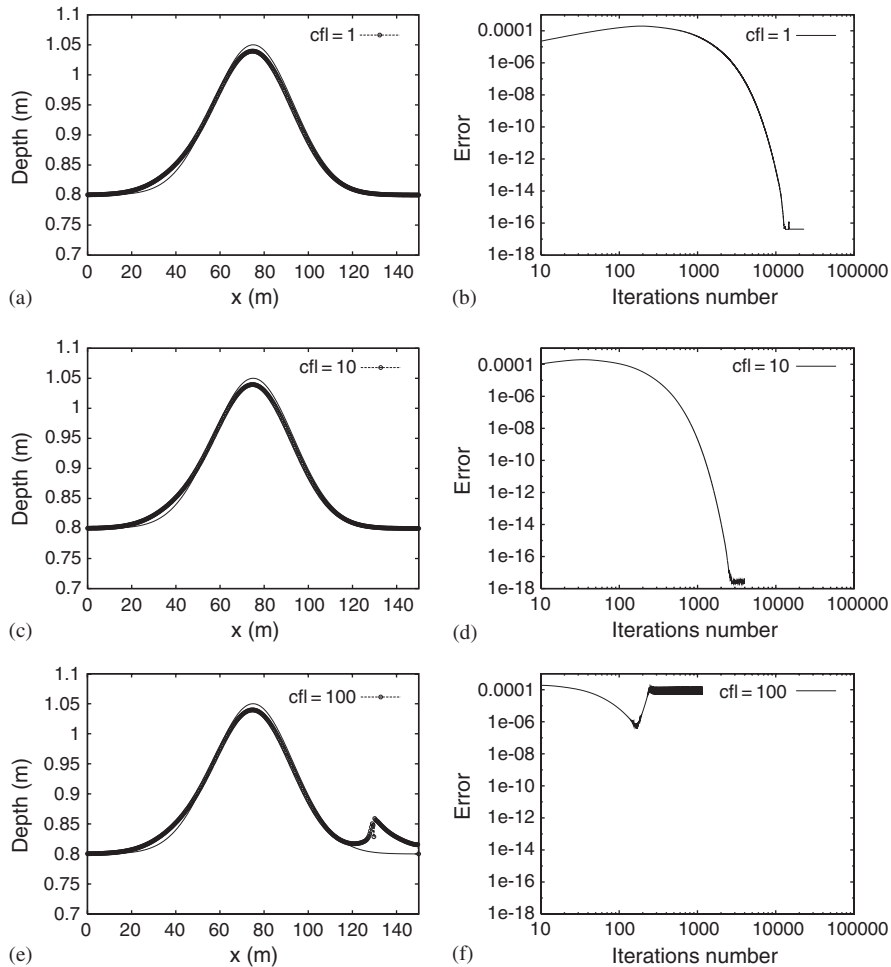


Figure 11. Numerical results of water depth and convergence error for the MacDonald 1 test case using the implicit bidiagonal scheme and CFL = 1, 10, 100.

affected by the discontinuous character of the initial conditions and are usually reported to fail for most applications using CFL values above 3 or 4. In the example presented here an initial depth ratio of 100 was used. The numerical solutions are displayed for comparison in Figure 8 together with the exact solution [20]. Figures 8(a) and (c) show the solution obtained with the bidiagonal implicit method using CFL = 1 and 2, respectively. The stability limit in this case for this scheme is shown in Figure 8(e) after one time step of calculation (approximately corresponding to CFL = 6). The oscillation shown makes a second time step impossible. On the contrary, Figures 8(b), (d) and (f) show the stability of the solution provided by the NDA semi-explicit scheme even using CFL = 100. It is worth to note that the solution is even more accurate at low CFL values in this case than using the bidiagonal implicit. The stability of the NDA scheme with CFL = 100 contrasts again with the strange

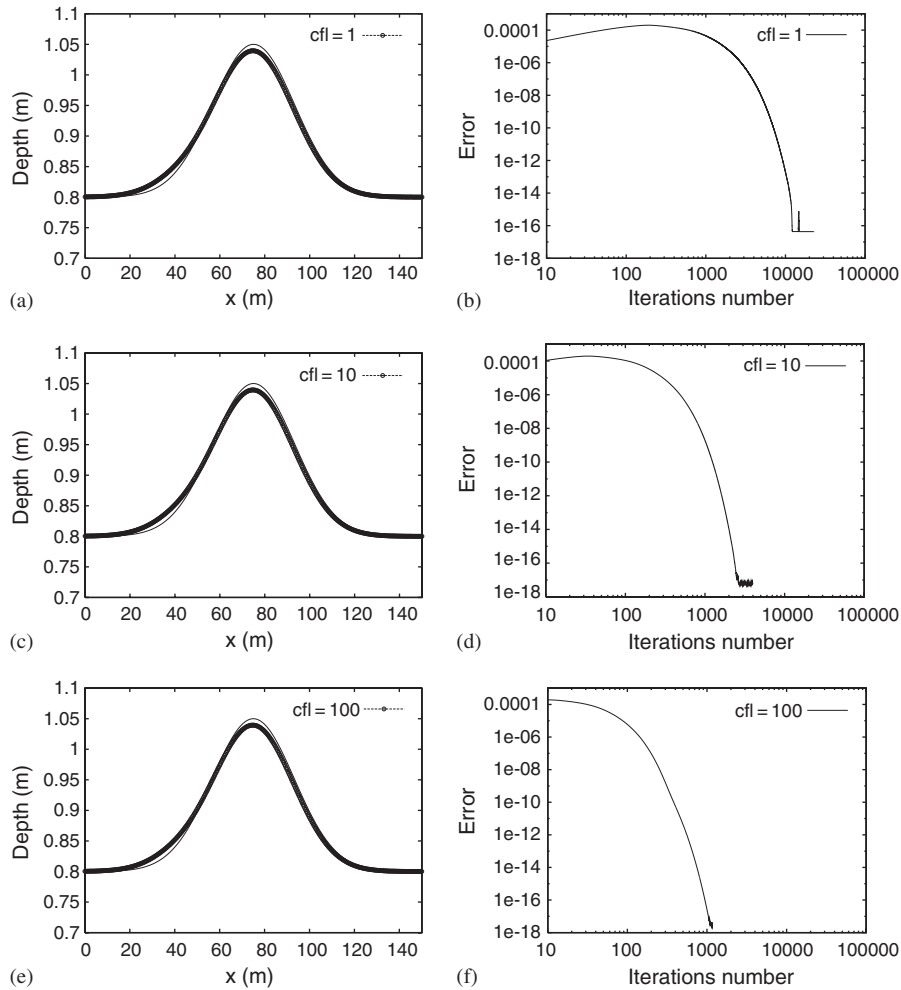


Figure 12. Numerical results of water depth and convergence error for the MacDonald 1 test case using the NDA semi-explicit scheme and CFL = 1, 10, 100.

shape of the solution at the first time steps. Figure 8(f) is the plot of the unsteady solution obtained in only a pair of time steps starting from discontinuous initial conditions. In order to follow how this solution evolves in time, Figure 9 shows the water depth at times $t = 60$ and 150 s in a longer domain.

7.3. Steady-state test cases

MacDonald *et al.* [21] 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. Two examples from Reference [22] are used here. They both apply a Manning friction parameter $n = 0.03$ and have been simulated using 400 cells. In the first example the flow of discharge

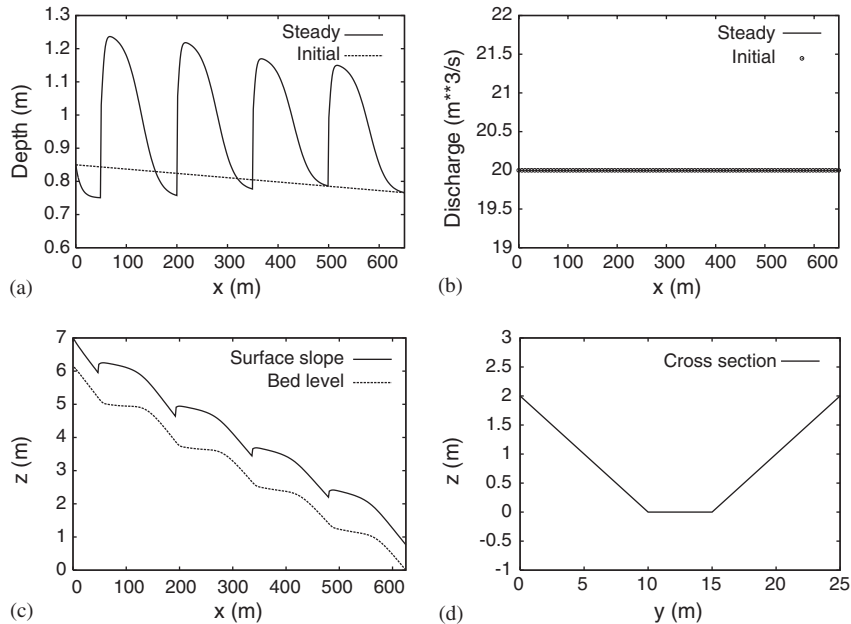


Figure 13. MacDonald 2 test case: (a) Initial and steady water depth profiles; (b) initial and steady discharge profiles; (c) water surface and channel bed profiles; and (d) channel cross-section.

$Q = 20 \text{ m}^3/\text{s}$ is subcritical all along the 10 m wide ($B = 10 \text{ m}$) rectangular channel shown in Figure 10 where the bed slope variation as well as the initial conditions used are also displayed. The steady-state water depth and bed slope analytical functions are [22]:

$$h(x) = 0.8 + 0.25 \exp\left(-33.75 \left(\frac{x}{150} - \frac{1}{2}\right)^2\right)$$

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

The performance of the discussed implicit schemes is displayed in Figures 11 and 12 by means of the numerical water depth versus the analytical solution and corresponding convergence to steady-state error using CFL = 1, 10 and 100. The way to evaluate the convergence from initial to steady state was the mean quadratic error in discharge defined as

$$E = \frac{1}{N} \sqrt{\frac{\sum_{i=1}^N [Q_i^n - Q(x_i)]^2}{\sum_{i=1}^N [Q(x_i)]^2}} \quad (47)$$

being N the number of nodes and $Q(x_i)$ the analytical steady discharge at node i . In order to compare the two methods in a case with steady discontinuities, the MacDonald 2 test case was used. Its main features are graphically described in Figure 13. The numerical results using CFL = 1 can be seen in Figure 14 as computed with the implicit bidiagonal. Both schemes

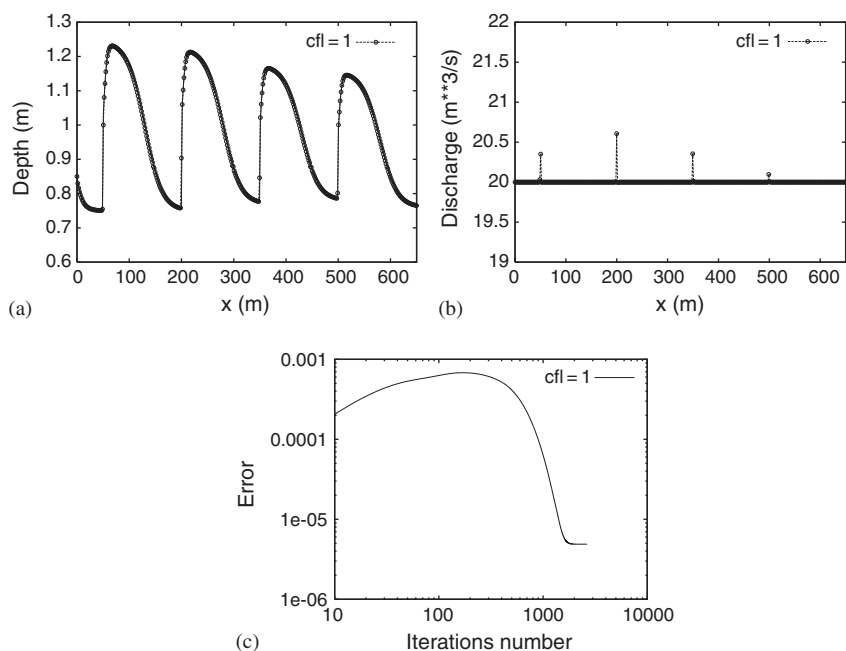


Figure 14. MacDonal 2 test case using bidiagonal implicit scheme with CFL = 1: (a) water depth profile; (b) discharge profile; and (c) convergence.

supply the same solution at this CFL number where the most noticeable is the local errors in discharge at the transcritical points. This can be seen in Figure 14. At higher CFL numbers there are more differences between the two solutions. Figure 15 shows the solution from the bidiagonal implicit using CFL = 10 and 25. The steady solution is very accurate in this case but CFL = 25 was the upper limit for the time step size compatible with stability. On the other hand, the results shown in Figure 16 indicate that there is not a stability limit when using the NDA semi-explicit and that a numerical diffusion, specially noticeable in the discharge plot Figure 16(c), appears due to the transcritical discontinuity and grows with the size of the time step. The solution is mass conservative in both cases, however, until the diffusion error interferes with the boundaries as in Figure 16(b,d).

7.4. Shallow water equations: flood wave in a river

The river reach used for the simulation belongs to the lower part of the Ebro river and therefore has very mild average slope and low water velocities. The river cross-section is highly variable in shape along the axis of the river and presents an irregular tendency in the bottom level variation leading to adverse and important slopes in some parts.

The total length of the simulated reach is around 11 400 m. Geometric data were available at 49 cross-sections. The Manning coefficient was assumed uniform and equal to 0.03.

A first run supplied the initial condition for the flooding simulation. This first run started from dry bed and introduced an upstream constant discharge of $Q = 200 \text{ m}^3/\text{s}$ until conver-

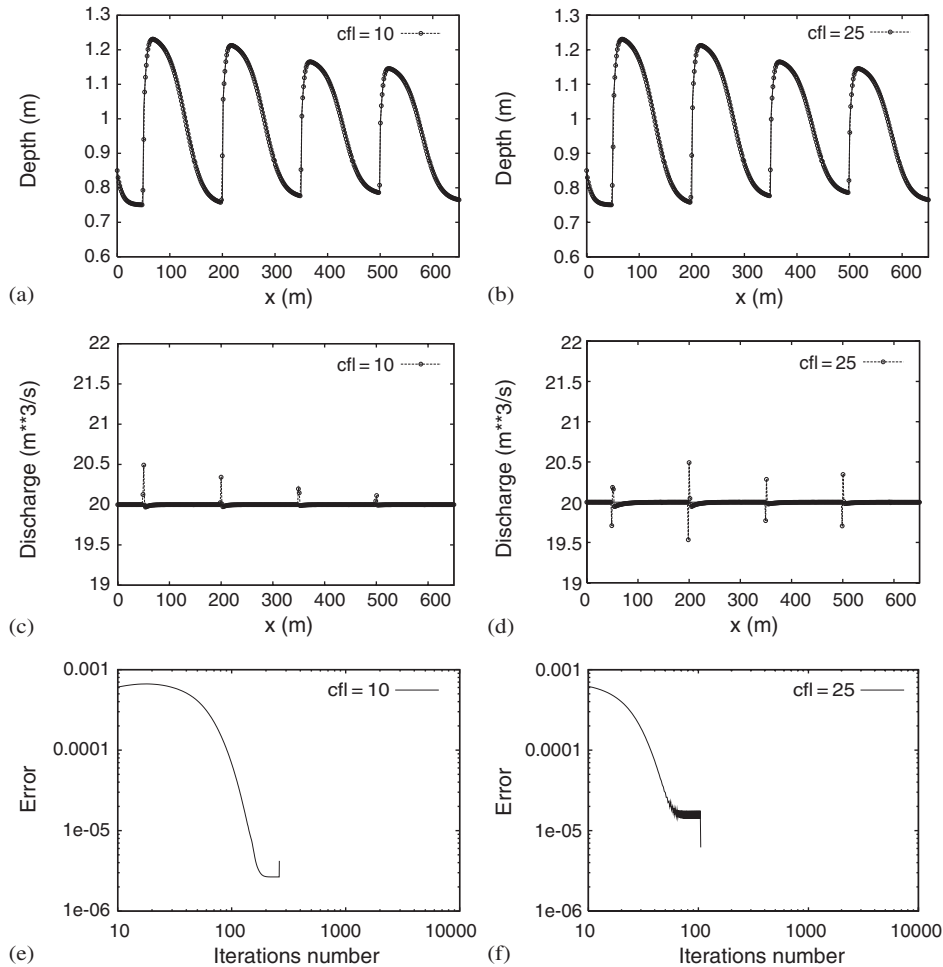


Figure 15. MacDonalD 2 test case. Bidiagonal implicit scheme.

gence. Unsteady shallow water flow over dry beds is at present one of the topics of research in computational hydraulics. Our approach consists of a through calculation of the front position as it advances over a computational mesh covering all the physical domain and in which there are both wet and dry cells. The set of equations we are solving are written in terms of wetted cross-section size and average velocity. The information concerning the original velocity profile is therefore reduced to a friction source term which takes into account the wall tangential stress and other viscous effects. This is usually modelled by means of a Manning-type formula which actually derives from the viscous nature of the velocity profile. In nature, rivers and streams do not have a regular bottom roughness, but the bed consists of unsorted sand, gravel and rocks. In cases of flow advance over dry bed, the flow resistance is mainly due to form drag and energy dissipation more than skin friction. Neither Chezy nor Manning formulae are adequate to estimate the friction losses and experimental investigations should be performed

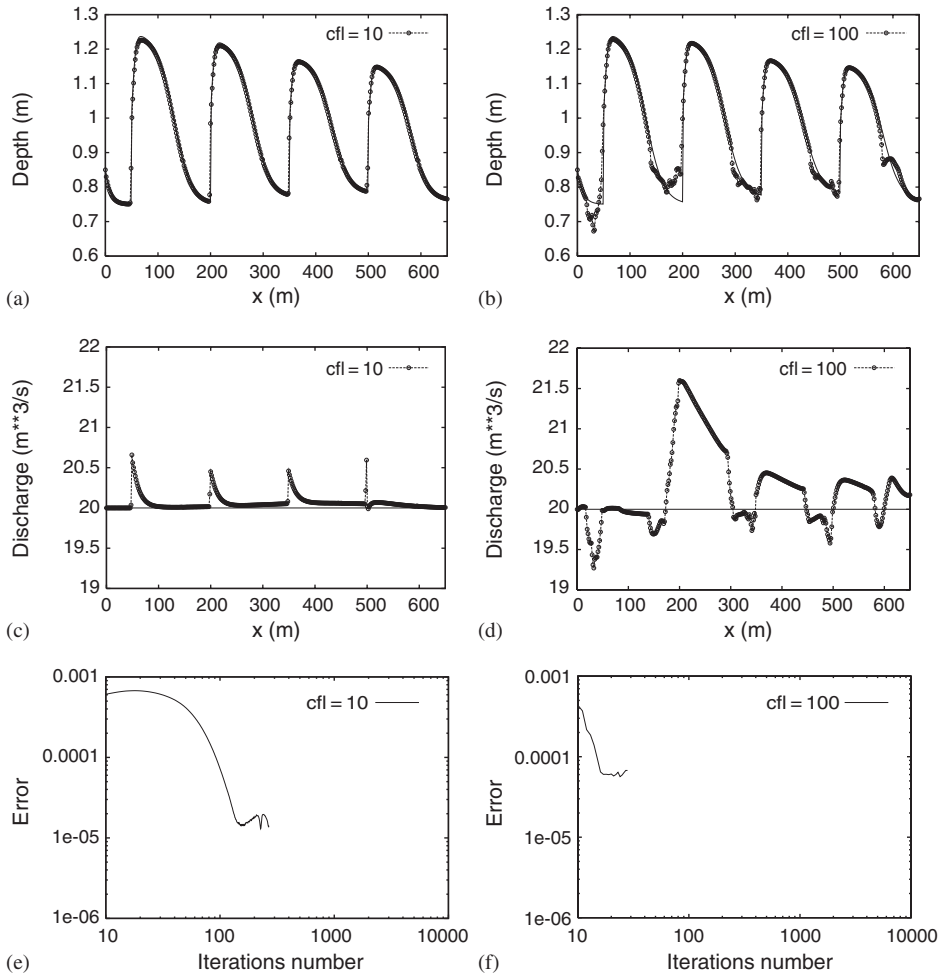


Figure 16. MacDonal 2 test case. NDA semi-implicit scheme.

beforehand. Owing to the impossibility of carrying out such field experiments in most cases a compromise is required. The option chosen in this work is based on the empirical correlation proposed by Strickler in 1923 for the Manning coefficient in rivers

$$n = 0.041d_{50}^{1/6}$$

Given an estimation for the global or local Manning coefficient, the above relation supplies the order of magnitude of d_{50} . In our model, this value is used as the minimum water depth required at the front position to allow front advance. For water depths below that value, water is forced to stop and accumulate.

Using the converged steady state as initial condition, the flood was represented by means of an upstream hydrograph, used as upstream boundary condition. The shape of this hydrograph

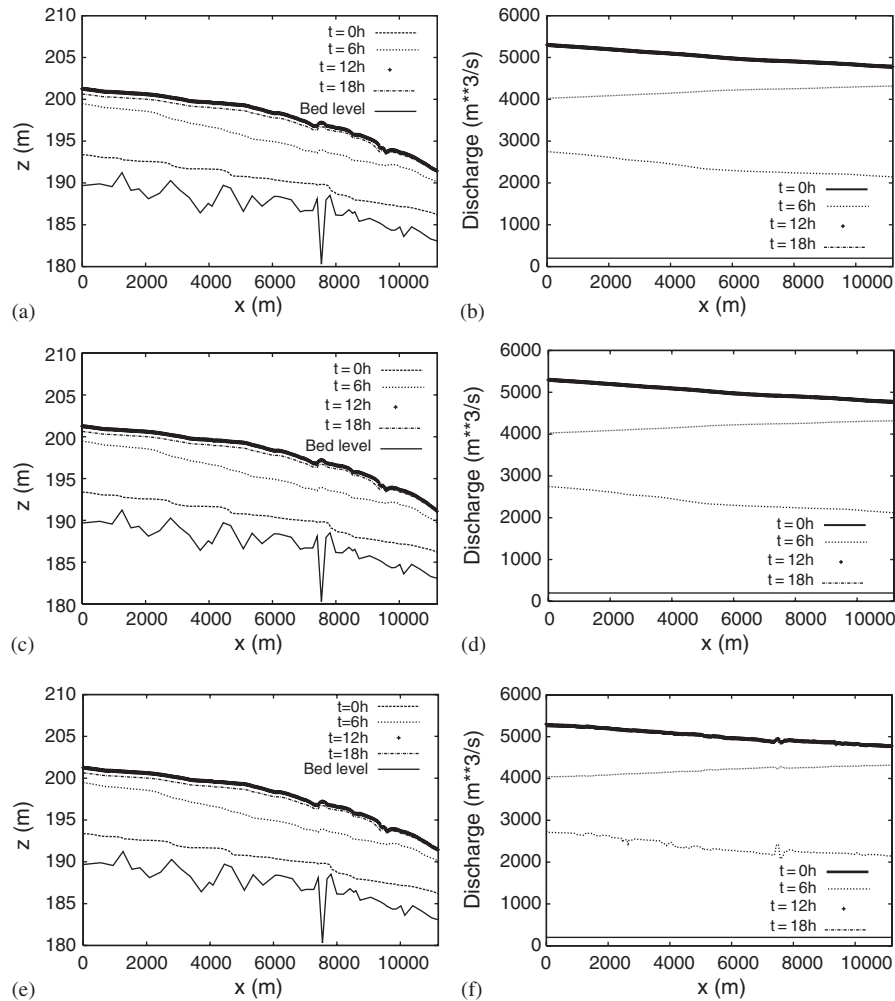


Figure 17. Water surface and discharge profiles at best flow ($t = 0$), $t = 6, 12$ h (peak inflow) and $t = 18$ h using (a, b) CFL = 1 (c, d), CFL = 10 and (e, f) CFL = 100 from the NDA scheme.

was simplified making it triangular and only the peak discharge, $Q = 5300 \text{ m}^3/\text{s}$, corresponded to the estimated maximum discharge for the flood event of return period equal to 500 years. The downstream boundary condition was set to critical flow in all cases.

From the numerical point of view, 400 nodal values were used and different values for the CFL number. Figure 17 shows the results obtained with the NDA semi-explicit scheme and CFL = 1, 10 and 100, leading to CPU times of around 90, 10 and 2 s, respectively, for the 36 h simulated event on a 1 GHz Pentium III PC. For the sake of comparison in this real test case, the same simulation was performed with the explicit first-order upwind and the bidiagonal implicit upwind schemes. The results are presented in Figure 18 using CFL = 1

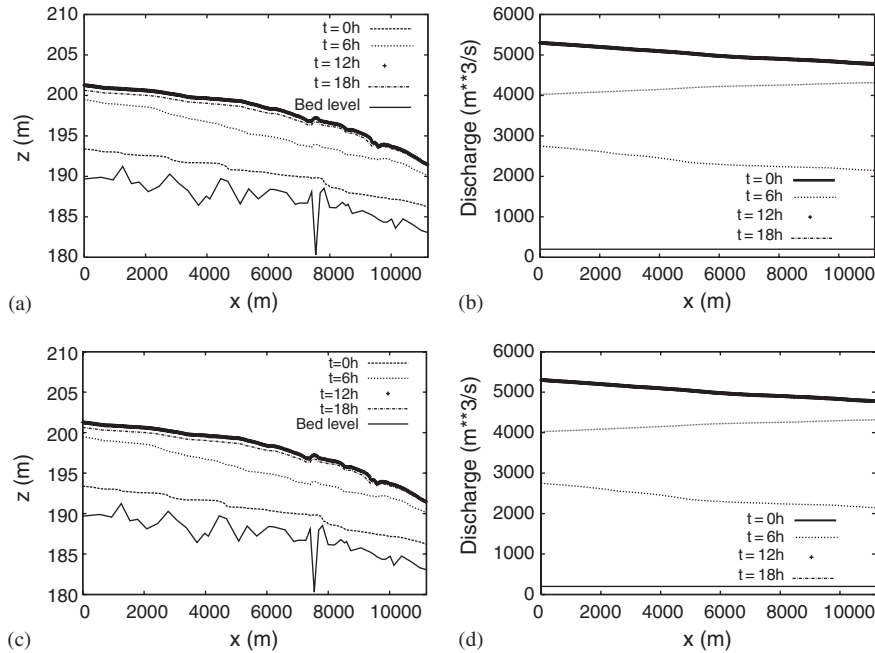


Figure 18. Water surface and discharge profiles at base flow ($t = 0$), $t = 6, 12$ h (peak inflow) and $t = 18$ h using (a, b) Roe's first-order explicit scheme with $CFL = 1$ and (c, d) the bidiagonal implicit scheme with $CFL = 5$.

(CPU time = 90 s) for the explicit scheme and $CFL = 5$ (CPU time = 10 s) for the bidiagonal implicit. Higher values of the CFL number were not possible for stable solutions using the bidiagonal scheme. The tridiagonal scheme is not even ready to compute flow on irregular geometries given its poor performance on idealized test cases.

The water depth and discharge profiles plotted and displayed correspond to the hypothesis of vertical walls at the left and right points of every supplied cross-section, that is, no lateral water losses are assumed. A control over the global volume conservation of the numerical results gives accuracy to machine precision in all cases. The results are remarkably similar and almost independent of the CFL (time step) used in this transient flow problem.

8. CONCLUSIONS

Conservative implicit methods belonging to the family of linearized upwind schemes have been presented. They are in general suitable for the numerical treatment of a variety of hydraulic problems. Among them, the basic linearized scheme is restricted to low CFL values for the simulation of unsteady transcritical flow with shocks. Two modifications of this scheme have been derived introducing the idea of splitting the scheme in a superposition of negative and positive wave influences: a bidiagonal implicit and a semi-explicit scheme.

The performance of the implicit schemes has been tested in two solutions of the Burgers' equation, in a dam-break problem arising from the shallow water equations, in two steady open channel flow problems and in a transient river flow example.

Both the bidiagonal implicit and the semi-explicit schemes are able to handle unsteady problems involving transcritical shocks. However, only the NDA semi-explicit can handle these problems with no stability restriction on the CFL at the cost of a numerical diffusion across the transcritical point that grows, in steady state, with the size of the time step. The reduction of the excess of diffusion keeping the unconditional stability is envisaged as further research work.

The conservative character of the schemes can get spoiled if the discretization at the boundaries is not treated with care. A technique that combines the external boundary conditions with the requirement of global mass conservation has been described, leading to conservative solutions up to machine accuracy.

APPENDIX A: STABILITY AND TOTAL VARIATION DIMINISHING PROPERTY

For a given linear advection equation

$$\frac{\partial u}{\partial t} + a \frac{\partial u}{\partial x} = 0$$

a general three-point implicit scheme adopts the following form:

$$\begin{aligned} u_i^{n+1} + A(\delta u_{i+1/2}^{n+1} + \delta u_{i-1/2}^{n+1}) + B(\delta u_{i+1/2}^{n+1} - \delta u_{i-1/2}^{n+1}) \\ = u_i^n + C(\delta u_{i+1/2}^n + \delta u_{i-1/2}^n) + D(\delta u_{i+1/2}^n - \delta u_{i-1/2}^n) \end{aligned} \quad (\text{A1})$$

The amplification factor G of any time-stepping scheme is defined as

$$G = \frac{u_i^{n+1}}{u_i^n} \quad (\text{A2})$$

In the general three-point scheme (A1), for every Fourier component e^{ikx} , the amplification factor is

$$G = \frac{1 + 4iC \sin \phi \cos \phi - 4D \sin^2 \phi}{1 + 4iA \sin \phi \cos \phi - 4B \sin^2 \phi} \quad (\text{A3})$$

with $\phi = \frac{1}{2} k \delta x$. A numerical scheme is stable for linear equations provided that

$$|G(\phi)| \leq 1 \quad (\text{A4})$$

Applying the inequality to (A1), the following conditions apply

$$2C^2 - D \leq 2A^2 - B, \quad 2D^2 - D \leq 2B^2 - B \quad (\text{A5})$$

A stronger restriction over the numerical solution is the property of TVD [9], defined to avoid any numerical oscillation in the solution. The total variation is defined as

$$TV^n = \sum_i |\delta u_{i+1/2}^n| \tag{A6}$$

and a scheme is said to be TVD whenever

$$TV^{n+1} \leq TV^n \tag{A7}$$

The sufficient (although not necessary) conditions ensuring that a centred three-point scheme as in (A1) applied to the scalar linear equation are [9]

$$B \leq -|A|, \quad \frac{1}{2} \geq D \geq |C| \tag{A8}$$

APPENDIX B: SPECIAL STABILITY CONDITIONS FOR IMPLICIT METHODS

Traditional techniques to deduce the stability or TVD conditions for a numerical scheme are usually based on a linear analysis, as stated before, and can prove insufficient and therefore unable to produce unconditionally stable methods in presence of discontinuous solutions to non-linear equations. An example of this kind of behaviour is shown in Figure B1, where the numerical and exact solutions to the inviscid Burgers' equation are plotted together. They correspond to the initial condition ($t=0$) and subsequent propagation ($t=40$ and 80) of a shock-type solution as computed using the previously presented tridiagonal first-order upwind scheme with $\theta=1$ and $CFL=4$. It can be observed that the solution is far from the unconditional TVD property.

In a search for numerical schemes well behaved in these cases, a new stability condition based on the non-amplification of discontinuities will be introduced next. A numerical scheme will be said to be NDA if the amplitude of the discontinuities present in homogeneous

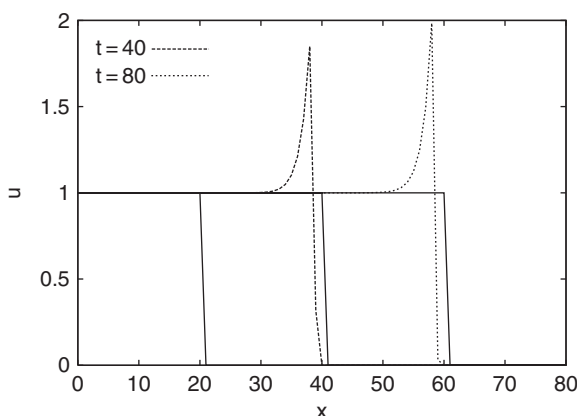


Figure B1. Numerical solution of a shock-type problem based on the Burgers' equation using the basic tridiagonal first-order upwind scheme ($CFL=4, \theta=1$ and $\delta x=1$).

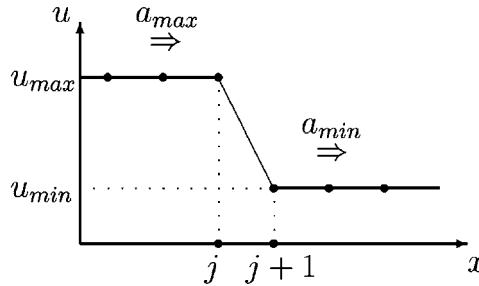


Figure B2. Initial conditions for a discontinuity propagation with a non-linear equation.

non-linear scalar equations does not grow in time due to the numerical propagation. Therefore, given a non-linear scalar equation like

$$\frac{\partial u}{\partial t} + a(u) \frac{\partial u}{\partial x} = 0 \tag{B1}$$

and discontinuous initial conditions as represented in Figure B2, we will study which conditions can guarantee that a given numerical scheme does not amplify the initial step.

It must be noted here that the limitation we want to impose may introduce some error in the discontinuity propagations since, in some cases, the spatial variation of the propagation velocity should require an increase on the discontinuity amplitude. However, we accept such error as necessary to ensure stability for the implicit methods in these cases using high values of the CFL number with non-linear equations.

It is possible to deduce the NDA conditions for the bidiagonal version of the first-order upwind implicit scheme. Let us assume, for the sake of simplicity, that the advection velocity is positive. Then for an equation like (B1), the bidiagonal scheme gives

$$\left[\left(1 + \theta \frac{\Delta t}{\delta x} a^+ \right) \Delta u \right]_i^n = \theta \frac{\Delta t}{\delta x} (a^+ \Delta u)_{i-1}^n - \frac{\Delta t}{\delta x} (a^+ \delta u)_{i-1/2}^n$$

With reference to Figure B2, assumed as the front position at time $t = n\Delta t$, the following can be stated:

For all points $i \leq j$, the step front has left, behind a uniform and constant state $u_i = u_{\max}$, hence,

$$\Delta u_i^n = \delta u_{i-1/2}^n = 0$$

When $i = j + 1$, $\Delta u_j^n = 0$, and

$$\left[\left(1 + \theta \frac{\Delta t}{\delta x} a^+ \right) \Delta u \right]_{j+1}^n = - \frac{\Delta t}{\delta x} (a^+ \delta u)_{j+1/2}^n$$

Using this result, it is not difficult to find that the NDA condition, formulated as

$$|\Delta u_{j+1}^n| \leq |\delta u_{j+1/2}^n|$$

leads to the following:

$$1 + \theta \frac{\Delta t}{\delta x} (a^+)_i^n \geq \frac{\Delta t}{\delta x} (a^+)_{i-1/2}^n \tag{B2}$$

On the other hand, for all points $i > j + 1$, and time $t = n\Delta t$, there is a uniform state $u_i = u_{\min}$, therefore

$$\begin{aligned} \delta u_{i-1/2}^n &= 0 \\ \left[\left(1 + \theta \frac{\Delta t}{\delta x} a^+ \right) \Delta u \right]_i^n &= \theta \frac{\Delta t}{\delta x} (a^+ \Delta u)_{i-1}^n \end{aligned}$$

Requiring, for the NDA behaviour, that

$$|\Delta u_i^n| \leq |\Delta u_{i-1}^n|$$

and leading to the additional condition

$$1 + \theta \frac{\Delta t}{\delta x} (a^+)_i^n \geq \theta \frac{\Delta t}{\delta x} (a^+)_{i-1}^n \tag{B3}$$

In the worst case, with $a_i = 0$, $a_{i-1} = a_{\max}$, $a_{i-1/2} = \frac{1}{2} a_{\max}$, condition (B2) indicates that the jump's amplitude will grow whenever the following holds:

$$1 \geq \frac{1}{2} \text{CFL}$$

and, therefore, the implicit upwind scheme applied to non-linear equations is only stable in presence of strong discontinuities for time steps such that $\text{CFL} \leq 2$.

REFERENCES

1. Preissmann A. Propagation des intumescences dans les canaux et rivières. First Congress of the French Association for Computation AFCAL, Grenoble, 1961; 443–442.
2. Abbott MB, Ionescu F. On the numerical computation of nearly horizontal flows. *Journal of Hydraulic Research* 1967; **5**(2):97–117.
3. Beam RM, Warming RF. An implicit factored scheme for the compressible Navier–Stokes equations. *AIAA Journal* 1978; **16**:393–402.
4. MacCormack RW. A numerical method for solving the equations of compressible viscous flow. *AIAA Journal* 1982; **20**:1275–1281.
5. Casier F, Deconinck H, Hirsch C. A class of bidiagonal schemes for solving Euler equations. *AIAA Journal* 1983; **22**:1556–1563.
6. Meselhe EA, Sotiropoulos F, Holly Jr FM. Numerical simulation of transcritical flow in open channels. *Journal of Hydraulic Engineering* 1997; **123**(9):774–782.
7. Yee HC. Construction of explicit and implicit symmetric TVD schemes and their applications. *Journal of Computational Physics* 1987; **68**:151–179.
8. Cunge JA, Holly FM, Verwey A. *Practical Aspects of Computational River Hydraulics*. Pitman: London, U.K., 1980.
9. Hirsch C. *Computational Methods for Inviscid and Viscous Flows: Numerical Computation of Internal and External Flows*. Wiley: New York, 1990.
10. 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.

11. Glaister P. Approximate Riemann solutions of the shallow water equations. *Journal of Hydraulic Research* 1988; **26**:293–306.
12. Bermúdez A, Vázquez Cendón ME. Upwind methods for hyperbolic conservation laws with source terms. *Computers and Fluids* 1994; **23**(8):1049–1071.
13. García Navarro P, Vázquez Cendón ME. Some considerations and improvements on the performance of Roe's schemes for 1D irregular geometries on numerical treatment of the source terms in the shallow water equations. *Computers and Fluids* 2000; **126**:26–40.
14. Roe PL. Approximate Riemann solvers, parameter vectors, and difference schemes. *Journal of Computational Physics* 1981; **43**(2):357–372.
15. Harten A, Hyman JM. Self adjusting grid methods for one-dimensional hyperbolic conservation laws. *Journal of Computational Physics* 1983; **50**:235–269.
16. Yee HC, Beam RM, Warming RF. Boundary approximations for implicit schemes for one-dimensional inviscid equations of gasdynamics. *AIAA Journal* 1982; **20**(9):1203–1211.
17. Burguete J, García Navarro P, Aliod R. Numerical simulation of runoff from extreme rainfall events in a mountain water catchment. *Natural Hazards in Earth System Sciences* 2002; **2**:1–9.
18. Burguete J. Modelos unidimensionales de flujos de superficie libre y transporte en geometrías irregulares: Aplicación al flujo en ríos. *Ph.D. Thesis*, University of Zaragoza.
19. Burgers JM. A mathematical model illustrating the theory of turbulence. *Advances in Applied Mechanics* 1948; **1**:171–199.
20. Stoker JJ. *Water Waves*. Interscience Publishers: New York, 1957.
21. MacDonald I, Baines MJ, Nichols NK, Samuels PG. Analytical benchmark solutions for open-channel flows. *ASCE Journal of Hydraulic Engineering* 1997; **123**(11):1041–1045.
22. MacDonald I. Analysis and computation of steady open channel flow. *Ph.D. Thesis*, University of Reading, 1996.