



Weak solutions for partial differential equations with source terms: Application to the shallow water equations

J. Murillo*, P. García-Navarro

Fluid Mechanics. C.P.S. University of Zaragoza. Zaragoza, Spain

ARTICLE INFO

Article history:

Received 6 November 2009

Received in revised form 5 February 2010

Accepted 16 February 2010

Available online 21 February 2010

MSC:

35L65

65M06

65M12

76M12

76M20

Keywords:

Weak solutions

Well-balanced approach

Roe methods

Energy dissipating solutions

Stability region

Wet/dry front

Shallow water systems

ABSTRACT

Weak solutions of problems with m equations with source terms are proposed using an augmented Riemann solver defined by $m + 1$ states instead of increasing the number of involved equations. These weak solutions use propagating jump discontinuities connecting the $m + 1$ states to approximate the Riemann solution. The average of the propagated waves in the computational cell leads to a reinterpretation of the Roe's approach and in the upwind treatment of the source term of Vázquez-Cendón. It is derived that the numerical scheme can not be formulated evaluating the physical flux function at the position of the initial discontinuities, as usually done in the homogeneous case. Positivity requirements over the values of the intermediate states are the only way to control the global stability of the method. Also it is shown that the definition of well-balanced equilibrium in trivial cases is not sufficient to provide correct results: it is necessary to provide discrete evaluations of the source term that ensure energy dissipating solutions when demanded. The one and two dimensional shallow water equations with source terms due to the bottom topography and friction are presented as case study. The stability region is shown to differ from the one defined for the case without source terms, and it can be derived that the appearance of negative values of the thickness of the water layer in the proximity of the wet/dry front is a particular case, of the wet/wet fronts. The consequence is a severe reduction in the magnitude of the allowable time step size if compared with the one obtained for the homogeneous case. Starting from this result, 1D and 2D numerical schemes are developed for both quadrilateral and triangular grids, enforcing conservation and positivity over the solution, allowing computationally efficient simulations by means of a reconstruction technique for the inner states of the weak solution that allows a recovery of the time step size.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

There is a wide range of physical situations, such as flow in open channels and rivers, tsunami and flood modeling, that can be mathematically represented by first-order non-linear systems of partial differential equations, whose derivation involves an assumption of the shallow water type. With rare exceptions, the governing equations are hyperbolic. The system of equations in realistic shallow water models include source terms, that is, terms that are functions of the vector of unknowns.

For sometime it has been accepted that the discretization of source terms can be as challenging as that of the non-linear advection terms. It must be said that for most cases, even naive discretizations of the source terms work reasonably well, but there are some well documented situations in which only sophisticated schemes can perform adequately. When solving real

* Corresponding author.

E-mail address: Javier.Murillo@unizar.es (J. Murillo).

problems one is likely to encounter all sorts of situations, with a high probability that naive schemes will compromise the quality and reliability of the solution.

When incorporating the presence of the source terms in a given specific finite volume scheme (Roe's scheme is used here) the main focus has been traditionally put on how the numerical scheme can be modified so that it maintains a discrete balance between flux and source terms. In the context of the discretization of hyperbolic systems of conservative laws a fundamental point has been to get schemes that satisfy the preservation of steady-states such as still water equilibrium in the context of the shallow water system. The difficulty to build such schemes was pointed out by several authors and led to the notion of well-balanced schemes [9,10,15,19–21].

In this framework the development of robust and efficient explicit finite volume models of shallow water flow has been the matter of recent research in the computational hydraulics literature. A few efforts have been reported on the search for the best methods able to preserve the exact conservation property (C-property) [20] in presence of flow over irregular geometries [4,11]. When dealing with simulation problems that involve bed variations and transient flow over a dry bed, these flow features impose a heavier restriction than the classical Courant–Friedrichs–Lewy (CFL) condition [6,9,19] on the time step size that may lead to inefficient computations. It is possible to avoid the necessity of reducing the time step and, at the same time, preventing instability and ensuring conservation at all times by a suitable flux difference redistribution [11,13].

It can be argued that the presence of source terms warrants the construction of new weak solutions appropriate to the nature of the equations, rather than the use of those constructed for the simple, homogeneous case. Even ensuring the discrete equilibrium formulated in well-balanced schemes, the direct application of the conclusions derived for the homogeneous case to cases with source terms leads to important difficulties. One of the most dramatic is the appearance of negative values of water depth, not only in wet/dry fronts, but also in initially wet/wet Riemann problems.

Gravity and friction are the main forces driving open channel flows. When using the shallow water model in hydraulic simulation, these forces participate in the dynamic equation as sources/sinks of momentum. In cases of steady shallow water flow with non-zero velocity, the discrete balance must be revisited [12].

George [7] presented a well balanced augmented approximate Riemann solver for the extended one dimensional shallow water equations including in the original solution vector two new variables: momentum flux and bottom surface. The solver is well-balanced and maintains a large class of steady states by the use of a properly defined steady state wave: a stationary jump discontinuity in the Riemann solution that acts as a source term. The idea of a stationary jump discontinuity is adapted to the method proposed in this work where the original system is not enlarged. Without modifying the original solution vector of conserved quantities (mass and momentum), we present augmented approximate Riemann solvers for the shallow water equations in the presence of a variable bottom surface and friction. They belong to the class of simple approximate solvers that use a set of propagating jump discontinuities, or waves, to approximate the true Riemann solution. Typically, a simple solver for a system of m conservation laws uses m such discontinuities. We present a three wave solver for the 1D shallow water equations system (two equations) and a four wave solver for the 2D case (three equations). In this work we go back to the original ideas of Roe using the upwind discretization of the source terms proposed by Vázquez-Cendón [20].

The outline of the paper is as follows: the one dimensional discretization is described first, for a scalar equation with source terms in Section 2 and then, in Section 3, for the 1D shallow water equations, followed by its application to a selection of results to show its accuracy. In Section 4 the generalization to two dimensions, illustrated using triangular grids, is presented and again applied to the shallow water equations. The final section contains the conclusions derived from the work.

2. One dimensional scalar Riemann problems with source terms

The basic ideas underlying this work can be illustrated by examining the nonlinear scalar equation,

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = s, \quad (1)$$

where $f(u)$ is a convex nonlinear function of u and s is a source term. From $f(u)$ it is possible to find an advection, or transport velocity λ :

$$\lambda = \frac{df}{du} \quad \lambda = \lambda(u). \quad (2)$$

We are interested in weak solutions of the Riemann Problem (hereafter RP) defined by the initial condition

$$u(x, 0) = \begin{cases} u_L & \text{if } x < 0, \\ u_R & \text{if } x > 0. \end{cases} \quad (3)$$

This case corresponds to a RP with a source function.

2.1. Numerical modelling

Weak solutions of the RP can be found using the integral form of (1). Given initial values u_L, u_R , a time interval $[0, 1]$ and a space interval $[-X, X]$, with X sufficiently large, $S < X$, with S the position of the fastest wave at $t = 1$, the exact solution u at time $t = 1$ satisfies the following conservation integral:

$$\int_{-X}^{+X} u(x, 1) dx = X(u_L + u_R) - (f(u_R) - f(u_L)) + \int_0^1 \int_{-X}^{+X} s dx dt. \tag{4}$$

In order to obtain a numerical solution of (1) the domain is divided in computational cells of constant size Δx : the interval of the i th cell is defined by $[x_{i-1/2}, x_{i+1/2}]$. Let Δt be the time step and $t^n = n\Delta t$ a generic time level. Assuming the usual notation we indicate with u_i^n the cell-average value of the solution $u(x, t)$ for the i th cell at time t^n :

$$u_i^n = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} u(x, t^n) dx, \tag{5}$$

u_i^n is therefore a piecewise approximation of the solution at time t^n .

The first order Godunov method, updates the averaged quantities one time-step assuming the following piecewise constant approximation,

$$u(x, 0) = \begin{cases} u_i & \text{if } x < 0, \\ u_{i+1} & \text{if } x > 0, \end{cases} \tag{6}$$

at each cell edge. The solution of the RP is evolved for a time equal to the time step and the resulting solution is cell-averaged again obtaining the piecewise solution at the new time level t^{n+1} .

In the Roe approach, the solution of each RP is obtained from the exact solution of a locally linearized problem defined by an approximate solution $\hat{u}(x, t)$. The approximate solution must fulfill the Consistency Condition [9] where the integral of the approximate solution $\hat{u}(x, t)$ of the linearized RP over a suitable control volume is equal to the integral of the exact solution over the same control volume. In our case the control volume is characterized by the cell size, $[-\frac{\Delta x}{2}, \frac{\Delta x}{2}]$, leading to

$$\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \hat{u}(x, 1) dx = \Delta x (u_{i+1} + u_i) - (f(u_{i+1}) - f(u_i)) + \int_0^1 \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} s dx dt. \tag{7}$$

Furthermore, as the source term s is not constant in time, we assume the following time linearization:

$$s_{i+1/2} = \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} s(x, 0) dx. \tag{8}$$

Now, the Roe's approximate solution is constructed defining the following linear RP:

$$\begin{aligned} \frac{\partial \hat{u}}{\partial t} + \lambda^*(u_{i+1}, u_i) \frac{\partial \hat{u}}{\partial x} &= 0, \\ \hat{u}(x, 0) &= \begin{cases} u_i & \text{if } x < 0, \\ u_{i+1} & \text{if } x > 0, \end{cases} \end{aligned} \tag{9}$$

where $\lambda^*(u_{i+1}, u_i)$ is a constant. Integrating over the same control volume as in (7)

$$\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \hat{u}(x, 1) dx = \Delta x (u_{i+1} + u_i) - \lambda_{i+1/2}^* (u_{i+1} - u_i). \tag{10}$$

Since we want to satisfy (7), the constraint that follows is:

$$(\delta f - s)_{i+1/2} = f(u_{i+1}) - f(u_i) - s_{i+1/2} = \lambda_{i+1/2}^* (u_{i+1} - u_i) \tag{11}$$

leading to the following equality

$$\lambda_{i+1/2}^* = \tilde{\lambda}_{i+1/2} \theta_{i+1/2}, \tag{12}$$

with

$$\tilde{\lambda}_{i+1/2} = \frac{f(u_{i+1}) - f(u_i)}{u_{i+1} - u_i}, \quad \theta_{i+1/2} = 1 - \frac{s_{i+1/2}}{f(u_{i+1}) - f(u_i)}, \tag{13}$$

where $\tilde{\lambda}$ is the advection velocity when the source term s is not present and θ is a measure of the relative importance of the source terms over the advective part.

2.1.1. A two wave approximate Riemann solution

A weak solution of the linear RP in (9) that satisfies (10) in the case $\tilde{\lambda}_{i+1/2} > 0$ is proposed:

$$\hat{u}(x, t) = \begin{cases} u_i & \text{if } x < 0, \\ u_{i+1}^{**} & \text{if } 0 < x < \tilde{\lambda}_{i+1/2} t, \\ u_{i+1} & \text{if } x > \tilde{\lambda}_{i+1/2} t, \end{cases} \tag{14}$$

where one wave is associated to the celerity $\tilde{\lambda}$ and the other wave is steady and with

$$u_{i+1}^{**} = u_{i+1} - (\theta \delta u)_{i+1/2} = u_i + \left(\frac{s}{\tilde{\lambda}}\right)_{i+1/2}. \tag{15}$$

Fig. 1 shows a sketch of the approximate solution in the particular case $\tilde{\lambda}_{i+1/2} > 0, u_{i+1} > u_i$ and $s_{i+1/2} < 0$.

In case that $\tilde{\lambda}_{i+1/2} < 0$ the proposed solution is:

$$\hat{u}(x, t) = \begin{cases} u_i & \text{if } x < \tilde{\lambda}_{i+1/2} t, \\ u_i^* & \text{if } \tilde{\lambda}_{i+1/2} t < x < 0, \\ u_{i+1} & \text{if } x > 0, \end{cases} \tag{16}$$

with

$$u_i^* = u_i + (\theta \delta u)_{i+1/2} = u_{i+1} - \left(\frac{s}{\tilde{\lambda}}\right)_{i+1/2}. \tag{17}$$

These solutions are then evolved for a time equal to the time step; the resulting solution is cell-averaged again obtaining the piecewise solution at the new time level t^{n+1} . Assuming that $\tilde{\lambda}_{i-1/2} > 0$ and $\tilde{\lambda}_{i+1/2} < 0$, the volume integral in the cell $[0, \Delta x] \times [0, \Delta t]$ represented in Fig. 2 is:

$$u_i^{n+1} \Delta x = u_i^{**} (\tilde{\lambda}_{i-1/2} \Delta t) + u_i^n (\Delta x - \tilde{\lambda}_{i-1/2} \Delta t + \tilde{\lambda}_{i+1/2} \Delta t) - u_i^* (\tilde{\lambda}_{i+1/2} \Delta t), \tag{18}$$

that can be rewritten as

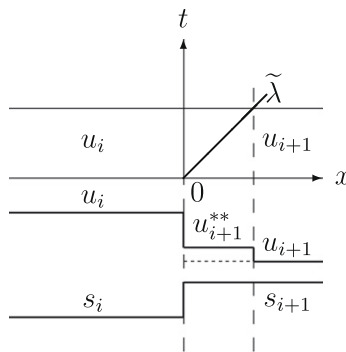


Fig. 1. Approximate solution for $\hat{u}(x, t)$.

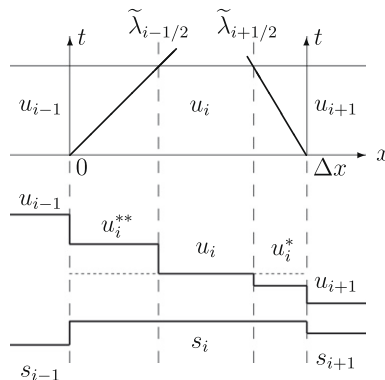


Fig. 2. Control volume in the Godunov method.

$$u_i^{n+1} = u_i^n - ((\delta u \theta \tilde{\lambda})_{i-1/2} + (\delta u \theta \tilde{\lambda})_{i+1/2}) \frac{\Delta t}{\Delta x}. \tag{19}$$

In Fig. 2 the time step is taken small enough so that there is no interaction of waves from neighboring Riemann problems. This would be necessary if we wanted to construct the solution at u_i^{n+1} in order to explicitly calculate the cell average (18). According to [9], in order to use the flux formula (19) it is only necessary that the edge value $\hat{u}(x, t)$ remains constant in time over the entire time step, which allows a time step roughly twice as large. Then, stability is ensured if the time step is limited by:

$$\Delta t \leq \Delta t^{\tilde{\lambda}}, \quad \Delta t^{\tilde{\lambda}} = \frac{\Delta x}{|\tilde{\lambda}|} \tag{20}$$

and the numerical scheme can be formulated in a general way as:

$$\begin{aligned} u_i^{n+1} &= u_i^n - ((\delta f - s)_{i-1/2}^+ + (\delta f - s)_{i+1/2}^-) \frac{\Delta t}{\Delta x}, \\ (\delta f - s)_{i+1/2}^\pm &= (\tilde{\lambda}^\pm \theta \delta u)_{i+1/2}. \end{aligned} \tag{21}$$

Straightforward algebraic manipulation converts (21) to an equivalent numerical flux-based finite volume scheme [8],

$$u_i^{n+1} = u_i^n - (f_{i+1/2}^* - f_{i-1/2}^*) \frac{\Delta t}{\Delta x} + (s_{i+1/2}^- + s_{i-1/2}^+) \frac{\Delta t}{\Delta x} \tag{22}$$

in which the numerical flux (denoted by an asterisk) for first-order upwinding is given by

$$f_{i+1/2}^* = \frac{1}{2} (f_{i+1} + f_i) - \frac{1}{2} (|\tilde{\lambda}| \delta u)_{i+1/2}, \tag{23}$$

with a similar expression for $f_{i-1/2}^*$, and

$$s_{i+1/2}^\pm = \frac{1}{2} (1 \pm \text{sgn}(\tilde{\lambda}))_{i+1/2} s_{i+1/2}. \tag{24}$$

These are essentially the two ways of extracting approximate information from the solution of the Riemann problem used in Godunov-type methods: one approach is, following Godunov’s original method, to use the Riemann solutions to determine cell intermediate fluxes at each time step as in (23); the other approach is based on the wave propagation algorithm in which waves arising in Riemann solutions are directly re-averaged onto adjacent cells in order to update the numerical solution as in (21). The second is applicable to equations in the form of a conservation law as well as those where there is not a flux function (non-conservative equations). Therefore, it is the latter route the one we follow in this work together with the new idea that the presence of the source term generates more than one intermediate state.

When $u_i^n \geq 0$ in all computational cells and positivity over the updated solutions is required, $u_i^{n+1} \geq 0$, the limit for the time step size in (20) is valid only if $u_i^* > 0$ and $u_{i+1}^{**} > 0$ is guaranteed. Otherwise, it is necessary to define additional restrictions over the cell average, isolating the contributions of each independent RP.

Then, with reference to Fig. 1, in the case $u_{i+1}^{**} \leq 0$ at a $i + 1/2$ edge, the time step must be redefined ensuring positivity on the cell average value in the control volume $[0, \frac{1}{2} \Delta x]$

$$(\tilde{\lambda}_{i+1/2} \Delta t) u_{i+1}^{**} + \left(\frac{1}{2} \Delta x - \tilde{\lambda}_{i+1/2} \Delta t \right) u_{i+1}^n \geq 0 \tag{25}$$

leading to the following limit in the size of the time step

$$\Delta t \leq \Delta t^{**}, \quad \Delta t^{**} = \frac{1}{2} \frac{\Delta x}{\tilde{\lambda}_{i+1/2}} \frac{u_{i+1}^n}{u_{i+1}^n - u_{i+1}^{**}}. \tag{26}$$

In case that $u_{i+1}^* \leq 0$ at a $i + 1/2$ edge, the limit is given by:

$$\Delta t \leq \Delta t^*, \quad \Delta t^* = \frac{1}{2} \frac{\Delta x}{|\tilde{\lambda}_{i+1/2}^1|} \frac{u_i^n}{u_i^n - u_i^*}. \tag{27}$$

It is worth noting that if $u_{i+1}^n = 0$ and $u_{i+1}^{**} < 0$ or if $u_i^n = 0$ and $u_i^* < 0$ the time step becomes nil according to (26) and to (27) respectively. Then, to ensure positivity of the solution in all cases the Godunov’s method for the scalar case is formulated as follows:

$$u_i^{n+1} = u_i^n - ((\delta f - s)_{i-1/2}^+ + (\delta f - s)_{i+1/2}^-) \frac{\Delta t}{\Delta x}, \tag{28}$$

where the fluxes in a general intercell edge $i + 1/2$ are computed as follows:

- If $u_{i+1}^n = 0$ and $u_{i+1}^{**} < 0$ set:

$$(\delta f - s)_{i+1/2}^- = 0, \quad (\delta f - s)_{i+1/2}^+ = 0. \quad (29)$$

- If $u_i^n = 0$ and $u_i^* < 0$ set:

$$(\delta f - s)_{i+1/2}^+ = 0 \quad (\delta f - s)_{i+1/2}^- = 0. \quad (30)$$

- Otherwise set:

$$(\delta f - s)_{i+1/2}^\pm = (\tilde{\lambda}^\pm \theta \delta u)_{i+1/2}, \quad (31)$$

with

$$\tilde{\lambda}_{i+1/2}^\pm = \frac{1}{2}(\tilde{\lambda} \pm |\tilde{\lambda}|). \quad (32)$$

In consequence the stability region becomes:

$$\Delta t \leq \min(\Delta t^{**}, \Delta t^*, \Delta t^{\tilde{\lambda}}), \quad (33)$$

where Δt^{**} is defined as in (26) if $u_{i+1}^{**} < 0$ and $u_{i+1}^n \neq 0$ and Δt^* is defined as in (27) if $u_i^* < 0$ and $u_i^n \neq 0$.

Depending on the particular approach chosen to discretize the integral source term $s_{i+1/2}$ the numerical scheme provides different solutions. The criterion to decide what is the best option for the source term discretization must be guided by the knowledge of the weak solution properties. This will be illustrated later with examples.

2.2. Application to the Burgers' equation with source terms

Consider Burgers' equation including source terms as proposed in [11]:

$$\frac{\partial u}{\partial t} + \frac{1}{2} \frac{\partial u^2}{\partial x} = -u \frac{\partial z}{\partial x}, \quad (34)$$

with the initial data

$$u(x, 0) = u_o(x) = \begin{cases} u_L & \text{if } x < 0, \\ u_R & \text{if } x > 0, \end{cases} \quad z(x) = \begin{cases} z_L & \text{if } x < 0, \\ z_R & \text{if } x > 0. \end{cases} \quad (35)$$

The exact solution to this problem is provided in Appendix A.

2.2.1. Numerical tests. Source term integration

In this section several RP solutions for Eq. (34) will be studied, analysing the evolution of the solution depending on the integration of the source term $s_{i+1/2}$ in (8). The cases are summarized in Table 1. All cases are computed using $\Delta x = 1$ and $CFL = 1$. Numerical solutions will be performed using numerical schemes (22) and (28) and the following approximate celerity $\tilde{\lambda}$:

$$\tilde{\lambda}_{i+1/2} = \frac{(u_{i+1} + u_i)}{2}. \quad (36)$$

Depending on the approach applied to integrate the discontinuous source term $s_{i+1/2}$ different solutions appear. Their properties will be stated from comparison with known exact solutions in order to select the best option. Three approaches for $s_{i+1/2}$ will be checked. The first approach for $s_{i+1/2}$, $s_{i+1/2}^a$, is constructed enforcing equilibrium in steady states:

$$s_{i+1/2}^a = -\frac{1}{2}(u_{i+1} + u_i)(z_{i+1} - z_i). \quad (37)$$

Two alternative approximations that do not enforce discrete equilibrium are also defined for numerical discussion:

Table 1
Summary of test cases.

Test case	u_L	u_R	z_L	z_R	$u_L \zeta$
1	2.0	1.0	0.0	0.5	1.5
2	2.0	1.0	0.0	-0.5	2.5
3	1.0	2.0	0.5	0.0	1.5
4	1.0	2.0	0.0	0.5	0.5
5	2.0	1.0	0.0	1.5	0.5
6	1.0	2.0	1.5	0.0	2.5
7	1.0	2.0	0.0	2.0	-1.0
8	2.0	0.0	0.0	2.5	-0.5

$$s_{i+1/2}^b = -u_{i+1}(z_{i+1} - z_i), \quad s_{i+1/2}^c = -u_i(z_{i+1} - z_i). \tag{38}$$

In test case 1 the exact solution is a right moving shock connecting u_R with $u^* = 1.5$ with a celerity $a = 1.25$. Fig. 3(a) shows the results given by the numerical scheme in (28) for approaches s^a, s^b and s^c . Function $s^a(-\circ-)$ provides an accurate value for u^* , while $s^b(-\bullet-)$ and $s^c(-\triangle-)$ overestimates and underestimates its value respectively, leading to a wrong shock velocity. In test case 2 the solution is a right moving shock connecting u_R with $u^* = 2.5$ with a celerity $a = 1.75$. Also the function s^a provides the best results, as depicted in Fig. 3(b). Approaches s^b and s^c lead to the same tendency as before.

Test cases 3 and 4 correspond to a right moving rarefaction wave. Fig. 4(a) and (b) shows the results at $t = 15$ given by the numerical scheme in (28) for test cases 3 and 4, respectively, using the integral approach s^a, s^b and s^c . Although the approximate solution is constructed assuming only jumps, the rarefaction is satisfactorily computed when using approach s^a in both cases, while $s^b(-\bullet-)$ and $s^c(-\triangle-)$ overestimates and underestimates the value of u_L^* respectively.

The good accordance among exact solutions and numerical results for approach s^a can be explained attending to the result given by the approximate linear obtained when using this option, as in in this case the value of $u_{i+1}^{*,a}$ in (14) becomes

$$u_{i+1}^{*,a} = u_i - \delta z \tag{39}$$

equal to the analytical value for u^* in the case of a shock wave (172) or equal to the analytical value for u_L^* in the case of a rarefaction (178).

In test case 5, $u_L > u_R$, and for the homogeneous case, the solution is a shock wave. The presence of the source term leads to a rarefaction characterized by a value of $u_L^* = 0.5$. Fig. 5(a) shows the results at $t = 15$ using numerical scheme in (28), and how approach s^a accurately reproduces the rarefaction wave. When using the approach s^c where the solution becomes a stable rarefaction, defined from an incorrect value of $u_L^* = u_{i+1}^{*,c} = 0$. If approach $s^b(-\circ-)$ is applied the initial solution becomes steady, as $u_{i+1} = u_{i+1}^{*,b}$. In absence of exact solutions this would lead to the misleading conclusion that the result given by approach s^b would be correct and the integral approach would be retained as a well balanced one.

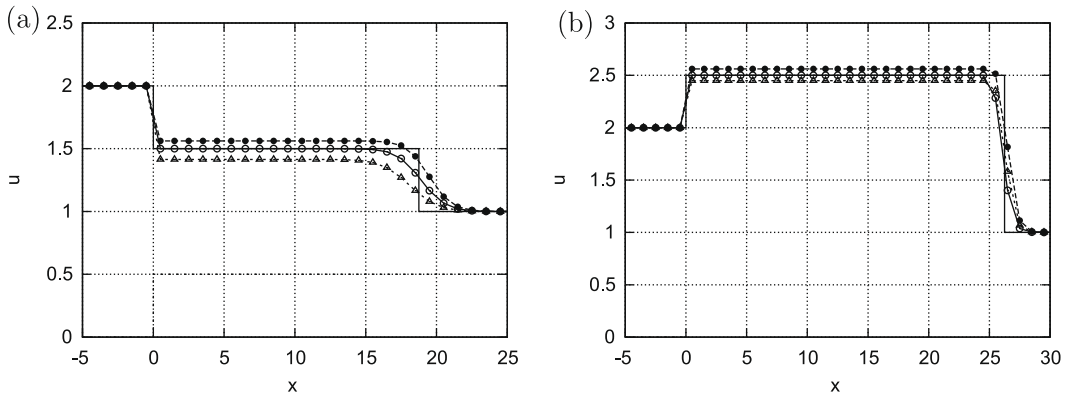


Fig. 3. (a) Exact solution (—) and computed solutions at $t = 15$ for test case 1, and (b) exact solution and computed solutions at $t = 15$ for test case 2, using approaches $s^a(-\circ-), s^b(-\bullet-)$ and $s^c(-\triangle-)$.

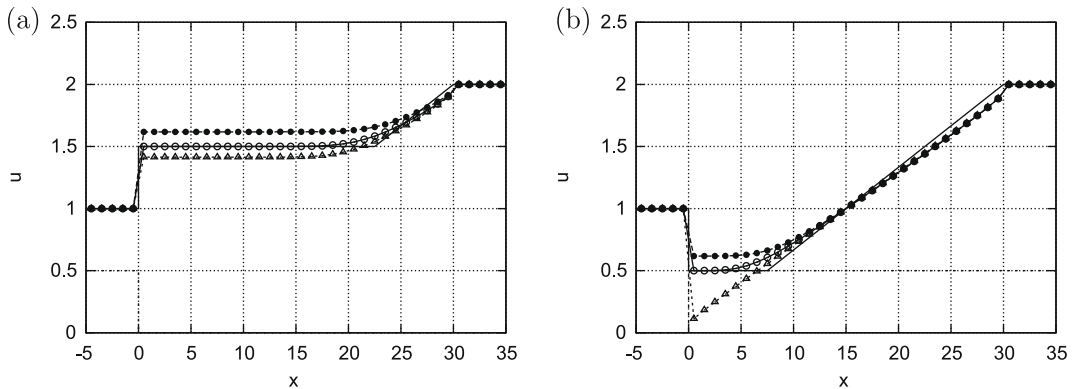


Fig. 4. (a) Exact solution (—) and computed solutions at $t = 15$ for test case 3, and (b) exact solution and computed solutions at $t = 15$ for test case 4, using approaches $s^a(-\circ-), s^b(-\bullet-)$ and $s^c(-\triangle-)$.

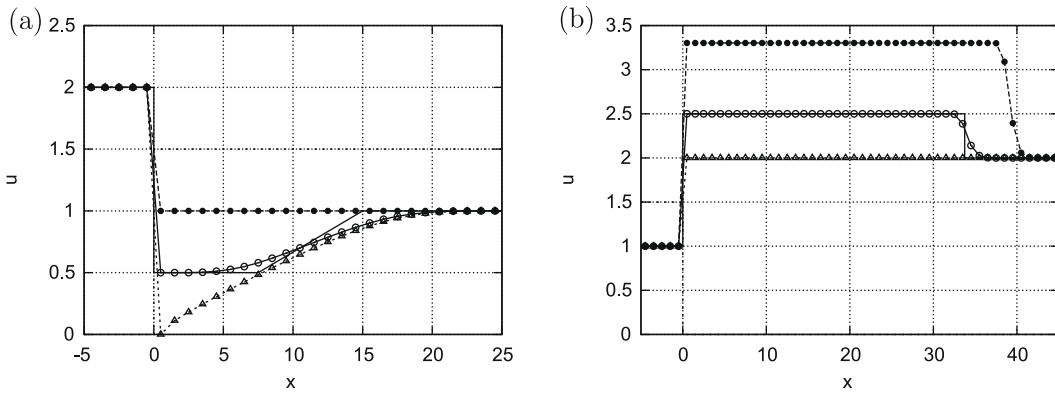


Fig. 5. (a) Exact solution (—) and computed solutions at $t = 15$ for test case 5, and (b) exact solution and computed solutions at $t = 15$ for test case 6, using approaches s^a (—○—), s^b (—●—) and s^c (—△—).

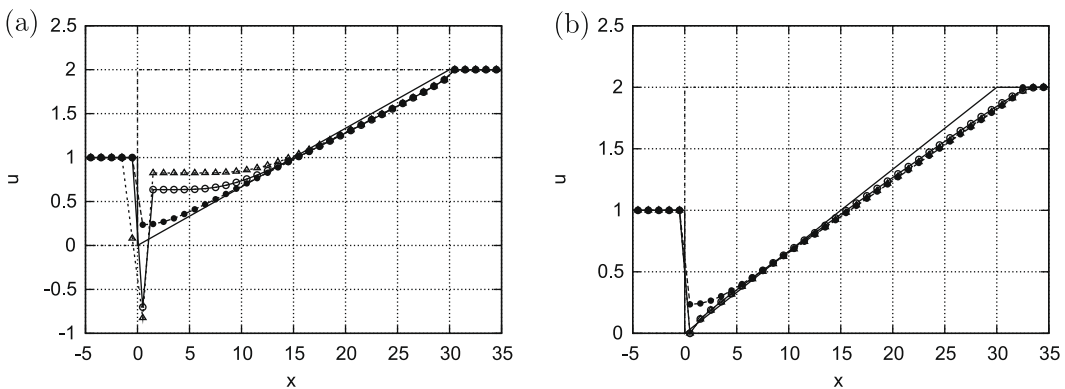


Fig. 6. (a) Exact solution (—) and computed solutions at $t = 15$ using numerical scheme (22), and (b) exact solution and computed solutions for numerical scheme (28), at $t = 15$ for test case 7, using approaches s^a (—○—), s^b (—●—) and s^c (—△—).

The solution for test case 6 for the homogeneous case is a rarefaction wave, but the presence of the source term leads to a shock wave, defined by $u^* = 2.5$ Fig. 5(b) shows the results at $t = 15$ for numerical scheme in (28). Integral approach s^a provides the most accurate solution while the integral approach s^b leads to an overestimation of u_L^* . If approach s^c is applied the initial solution becomes steady in time, as $u_{i+1} = u_{i+1}^{*,c}$. Again, in absence of analytical solutions, the result given by approach s^c may seem correct and would be retained as a well balanced approach.

In case 7, the initial conditions are defined to force null or negative values of $\hat{u}(x, t)$ in the right side. The numerical solution must converge asymptotically to a stable rarefaction between $u_L^* = 0$ and $u_R = 2$. When using numerical scheme in (22) with approach s_a , negative values of u appear, as depicted in Fig. 6(a). Only the solution for approach s_b remains positive, although an incorrect value of u_L^* is predicted. If, on the other hand, the scheme is carefully applied as in (28) the results improve as shown in Fig. 6(b). In this case both integral approaches s^a and s^c provide accurate and similar results, recovering the self-similarity of the problem. Approach s^b provides the same results as before.

For test case 8 the right state u_R is null, and the exact solution is a steady shock. If enforcing positivity over the solution, must remain steady and equal to the initial state. Fig. 7(a) shows how when using numerical scheme in (21) approaches s_a and s_c break the self similarity of the problem reaching negative values of u , while approach s_b leads to a result that may seem reasonable, a right moving wave, but incorrect. The results for the numerical scheme in (28) are depicted in Fig. 7(b), showing how both integral approaches s^a and s^c retain correctly the initial state, while integral approach s^b provides the same incorrect results.

3. 1D Systems of conservation laws with source terms

The discussion is next extended to hyperbolic nonlinear systems of equations with source terms in 1D, that expressed in integral formulation are:

$$\frac{\partial}{\partial t} \int_{x_1}^{x_2} \mathbf{U} dx + \mathbf{F}|_{x_2} - \mathbf{F}|_{x_1} - \int_{x_1}^{x_2} \mathbf{S} dx = 0, \tag{40}$$

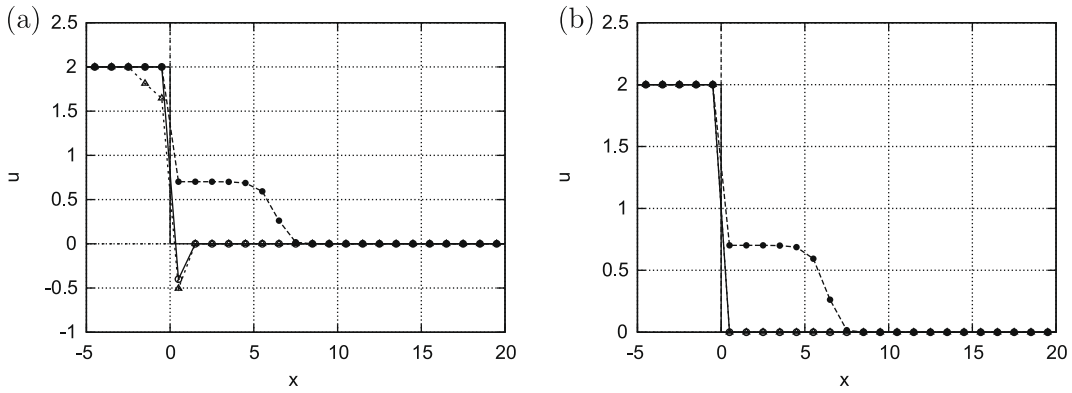


Fig. 7. (a) Exact solution (—) and computed solutions at $t = 15$ using numerical scheme (22), and (b) exact solution and computed solutions for numerical scheme (28), at $t = 15$ for test case 8, using approaches s^a (—○—), s^b (—■—) and s^c (—△—).

where x_1, x_2 are the limits of a generic control volume. The differential formulation is obtained assuming smooth variation of the variables and an infinitesimal width of the control volume:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = \mathbf{S} \tag{41}$$

and from this formulation it is possible to define a Jacobian matrix for the convective part \mathbf{J}

$$\mathbf{J} = \frac{d\mathbf{F}}{d\mathbf{U}}. \tag{42}$$

Assuming that the convective part in (40) is strictly hyperbolic with two real eigenvalues λ^1, λ^2 and eigenvectors $\mathbf{e}^1, \mathbf{e}^2$, it is possible to define two matrices $\mathbf{P} = (\mathbf{e}^1, \mathbf{e}^2)$ and \mathbf{P}^{-1} with the property that they diagonalize the Jacobian \mathbf{J}

$$\mathbf{J} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}. \tag{43}$$

3.1. Approximate solution of the Riemann problem

Given a Riemann problem for (41) with initial values $\mathbf{U}_L, \mathbf{U}_R$, a time interval $[0, 1]$ and a space interval $[-X, X]$, where

$$-X \leq \lambda_{min}, \quad X \geq \lambda_{max} \tag{44}$$

and $\lambda_{min}, \lambda_{max}$ are the positions of the slowest and the fastest wave at $t = 1$, the solution \mathbf{U} at time $t = 1$ satisfies the following property:

$$\int_{-X}^{+X} \mathbf{U}(x, 1) dx = X(\mathbf{U}_R + \mathbf{U}_L) - (\mathbf{F}(\mathbf{U}_R) - \mathbf{F}(\mathbf{U}_L)) + \int_0^1 \int_{-X}^{+X} \mathbf{S} dx dt. \tag{45}$$

In order to obtain a numerical solution of system (41) we divide the domain in computational cells of constant size Δx : the interval of the i th cell is defined by $[x_{i-1/2}, x_{i+1/2}]$ where $x_{i+1/2} = i\Delta x$ and the position of the center of the cell x_i is defined by $(i - 1/2)\Delta x$. Let Δt be the time step and $t^n = n\Delta t$ a generic time; assuming the usual notation we indicate with \mathbf{U}_i^n the cell-average value of the solution $\mathbf{U}(x, t)$ for the i th cell at time t^n :

$$\mathbf{U}_i^n = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{U}(x, t^n) dx. \tag{46}$$

\mathbf{U}_i^n is therefore a piecewise constant approximation of the solution at time t^n . The first order Godunov method, provides a way to update the averaged quantities one time-step in the following way: the piecewise approximations (46), are considered as initial values of local RPs:

$$\begin{aligned} \frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x}(\mathbf{U}_i, \mathbf{U}_{i+1}) &= \mathbf{S}(\mathbf{U}_i, \mathbf{U}_{i+1}), \\ \mathbf{U}(x, 0) &= \begin{cases} \mathbf{U}_i & \text{if } x < 0, \\ \mathbf{U}_{i+1} & \text{if } x > 0. \end{cases} \end{aligned} \tag{47}$$

These RP solutions are then evolved for a time equal to the time step; the resulting solution is cell-averaged again obtaining the piecewise solution at the new time level t^{n+1} . In the Roe approach, the solution of each RP is obtained from the exact solution of a locally linearized problem. This solution must fulfill the so called Consistency Condition, i.e. that the integral

of the solution $\widehat{\mathbf{U}}(x, t)$ of the linearized RP over a suitable control volume must be equal to the integral of the exact solution of (47) over the same control volume. Using (45) this condition becomes:

$$\int_{-X}^{+X} \widehat{\mathbf{U}}(x, 1) dx = X(\mathbf{U}_{i+1} + \mathbf{U}_i) - (\mathbf{F}(\mathbf{U}_{i+1}) - \mathbf{F}(\mathbf{U}_i)) + \int_0^1 \int_{-X}^{+X} \mathbf{S} dx dt. \tag{48}$$

Since the source term is not necessarily constant in time, we assume the following time linearization of the Consistency Condition:

$$\int_{-X}^{+X} \widehat{\mathbf{U}}(x, 1) dx = X(\mathbf{U}_{i+1} + \mathbf{U}_i) - (\mathbf{F}(\mathbf{U}_{i+1}) - \mathbf{F}(\mathbf{U}_i) - \mathbf{S}_{i+1/2}), \tag{49}$$

where

$$\mathbf{S}_{i+1/2} = \int_{-X}^{+X} \mathbf{S}(x, 0) dx \tag{50}$$

is a suitable numerical source vector.

In this formulation, RP (47) is approximated by using the following linear RP:

$$\begin{aligned} \frac{\partial \widehat{\mathbf{U}}}{\partial t} + \mathbf{J}^* \frac{\partial \widehat{\mathbf{U}}}{\partial x} &= 0, \\ \widehat{\mathbf{U}}(x, 0) &= \begin{cases} \mathbf{U}_i & \text{if } x < 0, \\ \mathbf{U}_{i+1} & \text{if } x > 0, \end{cases} \end{aligned} \tag{51}$$

where $\mathbf{J}^*(\mathbf{U}_i, \mathbf{U}_{i+1})$ is a constant matrix. Integrating (51) over the control volume $[-X, X] \times [0, 1]$, where X satisfies (44):

$$\int_{-X}^{+X} \widehat{\mathbf{U}}(x, 1) dx = X(\mathbf{U}_{i+1} + \mathbf{U}_i) - \mathbf{J}^*(\mathbf{U}_i, \mathbf{U}_{i+1}) (\mathbf{U}_{i+1} - \mathbf{U}_i). \tag{52}$$

Since we want to satisfy (49), the constraint that follows is:

$$\delta \mathbf{F}_{i+1/2} - \mathbf{S}_{i+1/2} = \mathbf{J}_{i+1/2}^* \delta \mathbf{U}_{i+1/2}, \tag{53}$$

with $\delta \mathbf{F}_{i+1/2} = \mathbf{F}(\mathbf{U}_{i+1}) - \mathbf{F}(\mathbf{U}_i)$ and $\delta \mathbf{U}_{i+1/2} = \mathbf{U}_{i+1} - \mathbf{U}_i$. Moreover, two more conditions are standard requirements for the Roe method.

$$\begin{aligned} \mathbf{J}_{i+1/2}^*(\mathbf{U}_{i+1}, \mathbf{U}_i), & \text{ is diagonalizable with real eigenvalues} \\ \mathbf{J}_{i+1/2}^*(\mathbf{U}_{i+1}, \mathbf{U}_i) & \rightarrow \mathbf{J}_{i+1/2}^*(\mathbf{U}_i) \text{ smoothly as } \mathbf{U}_{i+1} \rightarrow \mathbf{U}_i \end{aligned} \tag{54}$$

Considering that it is possible to define an approximate Jacobian $\widetilde{\mathbf{J}}_{i+1/2}$ for the homogeneous part, characterized by a set of approximate eigenvalues $\widetilde{\lambda}^1, \widetilde{\lambda}^2$ and eigenvectors $\widetilde{\mathbf{e}}^1, \widetilde{\mathbf{e}}^2$, two approximate matrices, $\widetilde{\mathbf{P}} = (\widetilde{\mathbf{e}}^1, \widetilde{\mathbf{e}}^2)$ and $\widetilde{\mathbf{P}}^{-1}$ are built with the following property:

$$\widetilde{\mathbf{J}}_{i+1/2} = \widetilde{\mathbf{P}}_{i+1/2} \widetilde{\mathbf{\Lambda}}_{i+1/2} \widetilde{\mathbf{P}}_{i+1/2}^{-1}. \tag{55}$$

The difference in vector \mathbf{U} across the grid edge is projected onto the matrix eigenvectors basis and the same for the source term:

$$\delta \mathbf{U}_{i+1/2} = \widetilde{\mathbf{P}}_{i+1/2} \mathbf{A}_{i+1/2}, \quad \mathbf{S}_{i+1/2} = \widetilde{\mathbf{P}}_{i+1/2} \mathbf{B}_{i+1/2} \tag{56}$$

with $\mathbf{A}_{i+1/2} = (\alpha^1 \quad \alpha^2)^T_{i+1/2}$ and $\mathbf{B}_{i+1/2} = (\beta^1 \quad \beta^2)^T_{i+1/2}$. Expressing all terms more compactly:

$$\delta \mathbf{F}_{i+1/2} - \mathbf{S}_{i+1/2} = \sum_{m=1}^{N_s} (\widetilde{\lambda}^{*m} \alpha \widetilde{\mathbf{e}})_{i+1/2}^m, \tag{57}$$

where

$$\widetilde{\lambda}_{i+1/2}^{*,m} = \widetilde{\lambda}_{i+1/2}^m \theta_{i+1/2}^m, \quad \theta_{i+1/2}^m = \left(1 - \frac{\beta}{\lambda \alpha} \right)_{i+1/2}^m, \tag{58}$$

so that the desired matrix in (52) is

$$\mathbf{J}_{i+1/2}^* = (\widetilde{\mathbf{P}} \widetilde{\mathbf{\Lambda}}^* \widetilde{\mathbf{P}}^{-1})_{i+1/2}, \tag{59}$$

with $\widetilde{\mathbf{\Lambda}}^* = \widetilde{\mathbf{\Lambda}} \Theta$, where $\widetilde{\mathbf{\Lambda}}_{i+1/2}$ is a diagonal matrix with eigenvalues $\widetilde{\lambda}_{i+1/2}^m$ in the main diagonal:

$$\widetilde{\mathbf{\Lambda}}_{i+1/2} = \begin{pmatrix} \widetilde{\lambda}^1 & 0 \\ 0 & \widetilde{\lambda}^2 \end{pmatrix}_{i+1/2} \tag{60}$$

and $\Theta_{i+1/2}$ is a diagonal matrix with $\theta_{i+1/2}^n$ in the main diagonal:

$$\Theta_{i+1/2} = \begin{pmatrix} \theta^1 & 0 \\ 0 & \theta^2 \end{pmatrix}_{i+1/2}. \tag{61}$$

3.2. Application to the 1D shallow water equations

For the 1D shallow water equations the relevant integral formulation in (40) derives from the depth-averaged equations of mass conservation and of momentum, with

$$\mathbf{U} = \begin{pmatrix} h \\ hu \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} 0 \\ \frac{p_b}{\rho_w} - \frac{\tau_b}{\rho_w} \end{pmatrix}, \tag{62}$$

where h represents the water depth, u the depth averaged component of the velocity vector and g is the acceleration of the gravity. The source term of the system is split in two kind of terms. The terms p_b and τ_b are the pressure along the bottom and the shear stress in the x direction respectively, with ρ_w the density of water. The above formulation is written in terms of the unit discharge and not valid for arbitrary cross sections. In order to extend the following discussion to general 1D problems [5] should be followed.

The convective part in (40) leads to two real eigenvalues λ^1, λ^2 and eigenvectors $\mathbf{e}^1, \mathbf{e}^2$,

$$\begin{aligned} \lambda^1 &= u - c, & \lambda^2 &= u + c, \\ \mathbf{e}^1 &= \begin{pmatrix} 1 \\ u - c \end{pmatrix}, & \mathbf{e}^2 &= \begin{pmatrix} 1 \\ u + c \end{pmatrix}, \end{aligned} \tag{63}$$

with $c = \sqrt{gh}$. Regarding the source term the following differential equation for the bottom slope can be obtained:

$$\frac{p_b}{\rho_w} = -gh \frac{\partial z}{\partial x} \tag{64}$$

The approximate Jacobian $\tilde{\mathbf{J}}$ for the homogeneous part [14] is

$$\tilde{\mathbf{J}}_{i+1/2} = \begin{pmatrix} 0 & 1 \\ \tilde{c}^2 - \tilde{u}^2 & 2\tilde{u} \end{pmatrix}_{i+1/2}, \quad \delta \mathbf{F}_{i+1/2} = \tilde{\mathbf{J}}_{i+1/2} \delta \mathbf{U}_{i+1/2} \tag{65}$$

with

$$\tilde{c} = \sqrt{g \frac{h_i + h_{i+1}}{2}}, \quad \tilde{u} = \frac{u_{i+1} \sqrt{h_{i+1}} + u_i \sqrt{h_i}}{\sqrt{h_{i+1}} + \sqrt{h_i}} \tag{66}$$

and the resulting set of approximate eigenvalues and eigenvectors are

$$\begin{aligned} \tilde{\lambda}^1 &= \tilde{u} - \tilde{c}, & \tilde{\lambda}^2 &= \tilde{u} + \tilde{c}, \\ \tilde{\mathbf{e}}^1 &= \begin{pmatrix} 1 \\ \tilde{u} - \tilde{c} \end{pmatrix}, & \tilde{\mathbf{e}}^2 &= \begin{pmatrix} 1 \\ \tilde{u} + \tilde{c} \end{pmatrix}. \end{aligned} \tag{67}$$

3.2.1. A three wave approximate Riemann solution

Depending on the flow conditions, three approximate solutions that satisfy (49), are proposed. The solutions for $\hat{\mathbf{U}}(x, t)$ are governed by the celerities in $\tilde{\Lambda}_{i+1/2}$ and each one consists of four regions.

The details of the approximate Riemann solution for each case are provided in Appendix B.

Following Godunov's method these RP solutions are then evolved for a time equal to the time step, the resulting solution is cell-averaged obtaining the piecewise solution at the new time level t^{n+1} . If both i and $i+1$ are subcritical, the integral volume in cell $[0, \Delta x] \times [0, \Delta t]$ is depicted in Fig. 8. Focusing on the updating rule for cell i :

$$\mathbf{U}_i^{n+1} \Delta x = \mathbf{U}_i^{n*} \left(\tilde{\lambda}_{i-1/2}^2 \Delta t \right) + \mathbf{U}_i^n \left(\Delta x - \tilde{\lambda}_{i-1/2}^2 \Delta t + \tilde{\lambda}_{i+1/2}^1 \Delta t \right) + \mathbf{U}_i^* \left(-\tilde{\lambda}_{i+1/2}^1 \Delta t \right), \tag{68}$$

that can be rewritten as

$$\mathbf{U}_i^{n+1} \Delta x = \mathbf{U}_i^n (\Delta x) + (\mathbf{U}_i^{n*} - \mathbf{U}_i^n) \left(\tilde{\lambda}_{i-1/2}^2 \Delta t \right) + (\mathbf{U}_i^n - \mathbf{U}_i^*) \left(\tilde{\lambda}_{i+1/2}^1 \Delta t \right) \tag{69}$$

and considering (183) the updated value \mathbf{U}_i^{n+1} is:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - (\theta \alpha \tilde{\mathbf{e}} \tilde{\lambda})_{i-1/2}^2 \frac{\Delta t}{\Delta x} - (\theta \alpha \tilde{\mathbf{e}} \tilde{\lambda})_{i+1/2}^1 \frac{\Delta t}{\Delta x}. \tag{70}$$

Straightforward algebraic manipulation converts (69) to an equivalent numerical flux-based finite volume scheme [8],

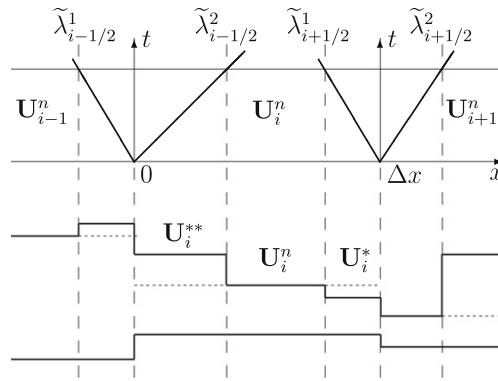


Fig. 8. Control volume in the Godunov method.

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \left(\mathbf{F}_{i+1/2}^* - \mathbf{F}_{i-1/2}^* \right) \frac{\Delta t}{\Delta x} + \left(\mathbf{S}_{i+1/2}^- + \mathbf{S}_{i-1/2}^+ \right) \frac{\Delta t}{\Delta x} \tag{71}$$

in which the numerical flux (denoted by an asterisk) for first-order upwinding is given by

$$\mathbf{F}_{i+1/2}^* = \frac{1}{2} (\mathbf{F}_{i+1} + \mathbf{F}_i) - \frac{1}{2} (\mathbf{P} |\tilde{\Lambda}| \mathbf{P}^{-1} \delta \mathbf{U})_{i+1/2}, \tag{72}$$

with a similar expression for $\mathbf{F}_{i-1/2}^*$ and

$$\mathbf{S}_{i+1/2}^\pm = (\mathbf{P} \mathbf{I}^\pm \mathbf{P}^{-1} \mathbf{S})_{i+1/2} \tag{73}$$

and $\mathbf{I}^\pm = \Lambda^{-1} \frac{1}{2} (\Lambda \pm |\Lambda|)$.

In Fig. 8 the time step is small enough so that there is no interaction of waves from neighboring Riemann problems. This would be necessary if we wanted to construct the solution at \mathbf{U}_i^{n+1} in order to explicitly calculate the cell average (68). If positivity of all water depth values in the solutions is guaranteed, $h_i^{**} \geq 0$ and $h_{i+1}^* \geq 0$, according to [9], in order to use the flux formula (70) it is only necessary that the edge values $\hat{\mathbf{U}}(x, t)$ remain constant in time over the entire time step, which allows a time step roughly twice as large and the time step is limited by

$$\Delta t \leq \Delta t^{\tilde{\lambda}}, \quad \Delta t^{\tilde{\lambda}} = \frac{\Delta x}{\max_{m=1,2} |\tilde{\lambda}^m|}. \tag{74}$$

As the cell average is constructed averaging with the terms \mathbf{U}_i^{**} and \mathbf{U}_{i+1}^* , the appearance of negative values of h_i^{**} and h_{i+1}^* must be considered. Fig. 9 represents a case with a negative value of h_{i+1}^{**} at a $i + 1/2$ edge, where the flow is locally subcritical as in (182). As the approximate solutions used in each RP are independent, it is necessary to define the time step ensuring that the cell average value in the control volume $[0, \frac{1}{2} \Delta x]$ remains positive

$$h_{i+1}^{n+1} \frac{1}{2} \Delta x = h_{i+1}^{**} (\tilde{\lambda}_{i+1/2}^2 \Delta t) + h_{i+1}^n \left(\frac{1}{2} \Delta x - \tilde{\lambda}_{i+1/2}^2 \right) \Delta t \geq 0 \tag{75}$$

leading to the following limit in the size of the time step

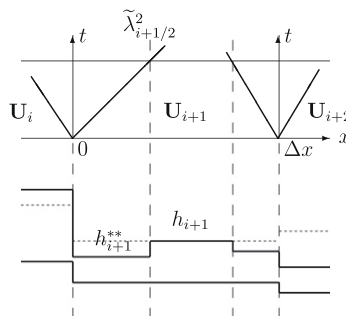


Fig. 9. Solution $\hat{\mathbf{U}}(x, t)$.

$$\Delta t \leq \Delta t^{**}, \quad \Delta t^{**} = \frac{1}{2} \frac{\Delta x}{\tilde{\lambda}_{i+1/2}^2} \frac{h_{i+1}^n}{h_{i+1}^n - h_{i+1}^{**}} \tag{76}$$

or in the case $h_i^* < 0$

$$\Delta t \leq \Delta t^*, \quad \Delta t^* = \frac{1}{2} \frac{\Delta x}{\tilde{\lambda}_{i+1/2}^1} \frac{h_i^n}{h_i^n - h_i^*}. \tag{77}$$

One case of special interest are wet/dry interfaces with discontinuous bed level, as it is possible to generate negative water depths in the initially dry region of $\widehat{\mathbf{u}}(\mathbf{x}, t)$. According to (76) or (77) the time step becomes nil in that case. To ensure positivity and conservation of the solution for all cases the Godunov’s method is formulated as follows:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \left((\delta \mathbf{F} - \mathbf{S})_{i-1/2}^+ + (\delta \mathbf{F} - \mathbf{S})_{i+1/2}^- \right) \frac{\Delta t}{\Delta x}, \tag{78}$$

where the fluxes in a general intercell edge $i + 1/2$ are computed as follows:

- If $h_{i+1}^n = 0$ and $h_{i+1}^{**} < 0$ set:

$$(\delta \mathbf{F} - \mathbf{S})_{i+1/2}^- = (\delta \mathbf{F} - \mathbf{S})_{i+1/2}, \quad (\delta \mathbf{F} - \mathbf{S})_{i+1/2}^+ = 0. \tag{79}$$

- If $h_i^n = 0$ and $h_i^* < 0$ set:

$$(\delta \mathbf{F} - \mathbf{S})_{i+1/2}^+ = (\delta \mathbf{F} - \mathbf{S})_{i+1/2}, \quad (\delta \mathbf{F} - \mathbf{S})_{i+1/2}^- = 0. \tag{80}$$

- Otherwise set:

$$(\delta \mathbf{F} - \mathbf{S})_{i+1/2}^\pm = \sum_{m=1}^{N_i} (\tilde{\lambda}^\pm \theta \alpha \tilde{\mathbf{e}})_{i+1/2}^m, \tag{81}$$

with

$$\tilde{\lambda}_{i+1/2}^{\pm m} = \frac{1}{2} (\tilde{\lambda} \pm |\tilde{\lambda}|). \tag{82}$$

The two main ways of formulating the Godunov type numerical solution for a system are then presented as it was done before for the scalar case. Note that because the numerical source integral cannot, in general, be written as a difference, it is not possible to include it in the numerical flux difference formulation (71). This means that the balance sought between flux derivatives and sources in the numerical flux based scheme can only be achieved locally by balancing non-zero fluxes through the edges of the control volume instead of setting every component to zero as in (78).

When in supercritical conditions values of $h_i^* < 0$ or $h_{i+1}^{**} < 0$ appear, the cell averaging in the Godunov method avoids negative values of h , as the source term does not participate in the updating of the water depth. In consequence, the stability region becomes:

$$\Delta t \leq \begin{cases} \min(\Delta t^{**}, \Delta t^*, \Delta t^{\tilde{\lambda}}) & \text{if } (\tilde{\lambda}^1 \tilde{\lambda}^2)_{i+1/2} < 0, \\ \Delta t^{\tilde{\lambda}} & \text{otherwise,} \end{cases} \tag{83}$$

where Δt^{**} is defined as in (76) if $h_{i+1}^{**} < 0$ and $h_{i+1}^n \neq 0$ and Δt^* is defined as in (77) if $h_i^* < 0$ and $h_i^n \neq 0$.

One result of Roe’s linearization is that the resulting approximate Riemann solution consists of only discontinuities and $\widehat{\mathbf{U}}(\mathbf{x}, t)$ is constructed as a sum of jumps or shocks. To avoid unphysical results the version of the Harten–Hyman entropy fix [18] is used. In the case of left transonic rarefaction $\lambda_i^1 < 0 < \lambda_{i+1}^1$, with $\lambda_i = \lambda(\mathbf{U}_i)$ and $\lambda_{i+1} = \lambda(\mathbf{U}_{i+1})$, the jump associated to $\tilde{\lambda}_{i+1/2}^1$ is decomposed into two new jumps,

$$\bar{\lambda}^1 = \lambda_i^1 \frac{(\lambda_{i+1}^1 - \tilde{\lambda}^1)}{(\lambda_{i+1}^1 - \lambda_i^1)}, \quad \hat{\lambda}^1 = \lambda_{i+1}^1 \frac{(\tilde{\lambda}^1 - \lambda_i^1)}{(\lambda_{i+1}^1 - \lambda_i^1)}, \tag{84}$$

with $\bar{\lambda}^1 + \hat{\lambda}^1 = \tilde{\lambda}^1$, and $\bar{\lambda}^1 < 0$ and $\hat{\lambda}^1 > 0$ by definition.

This idea can be applied to the decomposition of the source term associated to β^1 into two new values, $\bar{\beta}^1$ and $\hat{\beta}^1$. Their definition has to be done enforcing a conservative splitting of the source terms, $\bar{\beta}^1 + \hat{\beta}^1 = \beta^1$. Numerically it is possible to force a splitting proportional to the one performed on the $\tilde{\lambda}_{i+1/2}^1$ wave, that is

$$\bar{\beta}^1 = \beta^1 \frac{\bar{\lambda}^1}{\tilde{\lambda}^1}, \quad \hat{\beta}^1 = \beta^1 \frac{\hat{\lambda}^1}{\tilde{\lambda}^1} \tag{85}$$

but this option results in erroneous results and in a reduction of the time step size, as the values of $\bar{\beta}^1$ and $\hat{\beta}^1$ as defined in (85) are much greater than the original one, β^1 , as under these conditions $\tilde{\lambda}^1$ is a travelling wave with almost nil velocity. The option used in this work is

$$\bar{\beta}^1 = \beta^1, \quad \hat{\beta}^1 = 0, \tag{86}$$

that preserves the stability region in (83), simply replacing $\tilde{\lambda}_{i+1/2}^1$ by $\bar{\lambda}_{i+1/2}^1$.

For a right transonic rarefaction $\lambda_i^2 < 0 < \lambda_{i+1}^2$, the entropy fix procedure is entirely analogous to the left rarefaction case. The single jump in $\tilde{\lambda}^2$ is split into two smaller jumps $\bar{\lambda}^2$ and $\hat{\lambda}^2$

$$\bar{\lambda}^2 = \lambda_{i+1}^2 \frac{(\tilde{\lambda}^2 - \lambda_i^2)}{(\lambda_{i+1}^2 - \lambda_i^2)}, \quad \hat{\lambda}^2 = \lambda_i^2 \frac{(\lambda_{i+1}^2 - \tilde{\lambda}^2)}{(\lambda_{i+1}^2 - \lambda_i^2)}, \tag{87}$$

with $\bar{\lambda}^2 > 0$ and $\hat{\lambda}^2 < 0$ by definition. The source term is split enforcing

$$\beta^2 = \beta^2, \quad \bar{\beta}^2 = 0, \tag{88}$$

so the stability region in (83) is preserved, simply replacing $\tilde{\lambda}_{i+1/2}^2$ by $\bar{\lambda}_{i+1/2}^2$.

3.3. Integration of the source term

The source term in (50) is expressed as

$$\mathbf{S}_{i+1/2} = \begin{pmatrix} 0 \\ \frac{p_b}{\rho_w} - \frac{\tau_b}{\rho_w} \end{pmatrix}_{i+1/2}, \tag{89}$$

where $\frac{p_b}{\rho_w}$ and $\frac{\tau_b}{\rho_w}$ attend to the pressure and friction exerted on the bed respectively.

There is not a unique way to perform the numerical integral of the source term in (50). Under the hypothesis of smooth variation of the variables and an infinitesimal width of the control volume, it is possible to define the integral in (50) evaluating $\frac{p_b}{\rho_w}$ as

$$\left(\frac{p_b}{\rho_w}\right)_{i+1/2}^a = -g(\tilde{h}\delta z)_{i+1/2}, \tag{90}$$

with $\tilde{h} = 1/2(h_i + h_{i+1})$. Assuming a piecewise representation of the bed level, another possibility is to use the physical definition of the hydrostatic force exerted over the bed discontinuity so the pressure head depends only on the free-surface level. Attending to this definition, (50) is defined using the following approach for $\frac{p_b}{\rho_w}$:

$$\left(\frac{p_b}{\rho_w}\right)_{i+1/2}^b = -g\left(h_j - \frac{|\delta z'|}{2}\right)\delta z', \tag{91}$$

with

$$j = \begin{cases} i & \text{if } \delta z \geq 0, \\ i+1 & \text{if } \delta z < 0, \end{cases} \quad \delta z' = \begin{cases} h_i & \text{if } \delta z \geq 0 \text{ and } d_i < z_{i+1}, \\ h_{i+1} & \text{if } \delta z < 0 \text{ and } d_{i+1} < z_i, \\ \delta z & \text{otherwise,} \end{cases} \tag{92}$$

where $d = (h + z)$. Both approaches can be blended to provide another expression for the thrust term, that we will refer to as $\left(\frac{p_b}{\rho_w}\right)^c$:

$$\left(\frac{p_b}{\rho_w}\right)_{i+1/2}^c = \begin{cases} \max\left(\left(\frac{p_b}{\rho_w}\right)^a, \left(\frac{p_b}{\rho_w}\right)^b\right)_{i+1/2} & \text{if } \delta d \delta z \geq 0 \text{ and } \tilde{u}\delta z > 0, \\ \left(\frac{p_b}{\rho_w}\right)^b_{i+1/2} & \text{otherwise,} \end{cases} \tag{93}$$

that considers the problems associated to flow across an upward step in overtopping waves.

In cases of still water with a continuous water level surface all three approximations of the trust term, a , b and c , provide correct solutions for all values when constructing the approximate solution $\hat{\mathbf{U}}(x, t)$, as in this particular case, Fig. 10:

$$\begin{aligned} h_i^n + z_i &= h_i^* + z_i = h_{i+1}^{**} + z_{i+1} = h_{i+1}^n + z_{i+1}, \\ (hu)_i^n &= (hu)_i^* = (hu)_{i+1}^{**} = (hu)_{i+1}^n = 0. \end{aligned} \tag{94}$$

This is a particular case, and in cases of nonzero velocity, differences among the solutions provided by the source term integral in $(p_b/\rho_w)^a$, $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$ arise. One consequence of utmost importance is that they can generate negative values of water depth in the inner regions of the weak solution. In wetting/drying fronts negative values can be avoided if in each $i + 1/2$ edge with discontinuous water level surface, characterized by

$$h_i^n + z_i < z_{i+1}, \quad h_{i+1}^n = 0 \tag{95}$$

or

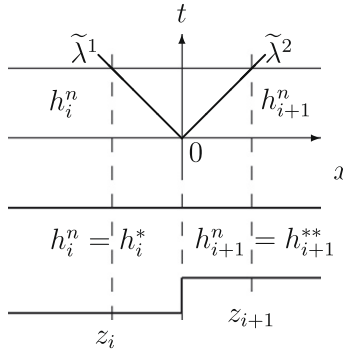


Fig. 10. Solution $\hat{\mathbf{U}}(x, t)$ in case of static equilibrium.

$$h_{i+1}^n + z_{i+1} < z_i, \quad h_i^n = 0, \tag{96}$$

a zero velocity in the involved cells is enforced, $u_{i+1}^n = u_i^n = 0$, in combination with approach $(p_b/\rho_w)^c$. This procedure, that considers the cell edges characterized by (95) or (96) as solid walls, will be referred to as approach $(p_b/\rho_w)^{c,w}$.

Regarding the friction term, the discretization based on [12] is applied

$$\frac{\tau_b}{\rho_w f_{i+1/2}} = g(\tilde{h}S_f)_{i+1/2}\Delta x, \quad S_{f,i+1/2} = \left(\frac{n^2 \tilde{u} |\tilde{u}|}{\max(h_i, h_{i+1})^{4/3}} \right)_{i+1/2}. \tag{97}$$

The importance of the discrete equilibrium in cases of still water to provide well-balanced schemes has been widely reported, but it is important to stress that with the unified formulation of the source terms, the scheme becomes well balanced in steady cases with no null velocity. According to numerical scheme (78), the stationary solution is reached when all updating components of the linearized solution become nil, that is $(\theta\alpha)_{i+1/2}^m = 0$ for $m = 1, 2$, leading to a constant discharge in all zones of the weak solution. In the subcritical case:

$$(hu)_i^n = (hu)_i^* = (hu)_{i+1}^{**} = (hu)_{i+1}^n \neq 0, \tag{98}$$

or in the supercritical case, with $u > 0$:

$$(hu)_i^n = (hu)_{i+1}^* = (hu)_{i+1}^{**} = (hu)_{i+1}^n \neq 0. \tag{99}$$

3.4. Reconstruction of the approximate solution $\hat{\mathbf{U}}(x, t)$

The linearization of the source terms leads to extremely small values of the allowable time step, as Δt^{**} or Δt^* can be various order of magnitude smaller than Δt^2 . This can be avoided by means of a reconstruction of the approximate solution $\hat{\mathbf{U}}(x, t)$. The strategy proposed here is based on enforcing positive values of h_i^* and h_{i+1}^{**} when they become negative. Considering that $\tilde{e}_1^1 = 1$, positive values of h_i^* require that

$$h_i^* = h_i^n + \alpha_{i+1/2}^1 - \left(\frac{\beta}{\tilde{\lambda}} \right)_{i+1/2}^1 \geq 0 \tag{100}$$

leading to the following limit over β^1

$$\beta_{i+1/2}^1 \geq \beta_{min}^1, \quad \beta_{min}^1 = -\left(h_i^n + \alpha_{i+1/2}^1 \right) |\tilde{\lambda}_{i+1/2}^1|. \tag{101}$$

Considering that $\tilde{e}_1^2 = 1$, positive values of h_{i+1}^{**} require that

$$h_{i+1}^{**} = h_{i+1}^n - \alpha_{i+1/2}^2 + \left(\frac{\beta}{\tilde{\lambda}} \right)_{i+1/2}^2 \geq 0 \tag{102}$$

and a limit over β^2 appears

$$\beta_{i+1/2}^2 \geq \beta_{min}^2, \quad \beta_{min}^2 = -\left(h_{i+1}^n - (\alpha)_{i+1/2}^2 \right) \tilde{\lambda}_{i+1/2}^2. \tag{103}$$

The reconstruction proposed in this work will be applied only to subcritical wet/wet RP, as in dry/wet RP the appearance of negative values of h_i^* or h_{i+1}^{**} in the approximate solution is helpful to provide a correct tracking of the flooding advance, and in supercritical cases the cell averaging of the weak solutions ensures positivity of the solution. Also cases where both $h_i^* < 0$ and $h_{i+1}^{**} < 0$ are omitted.

In the case $h_i^* < 0$ and $h_{i+1}^{**} > 0$ and $\Delta t^* < \Delta t^{\bar{z}}$, the new value of β^1 is redefined ensuring that h_{i+1}^{**} remains positive or null. To ensure conservation β^2 must be replaced by the new value of $-\beta^1$, then

$$\beta^1 = \begin{cases} \beta_{min}^1 & \text{if } -\beta_{min}^1 \geq \beta_{min}^2, \\ \beta^1 & \text{otherwise} \end{cases}, \quad \beta^2 = -\beta^1. \tag{104}$$

In the case $h_i^* > 0$ and $h_{i+1}^{**} < 0$ and $\Delta t^{**} < \Delta t^{\bar{z}}$, the new value of β^2 is redefined ensuring that h_i^* remains positive or null. To ensure conservation β^1 must be replaced by the new value of $-\beta^2$, then

$$\beta^2 = \begin{cases} \beta_{min}^2 & \text{if } -\beta_{min}^2 \geq \beta_{min}^1, \\ \beta^2 & \text{otherwise,} \end{cases} \quad \beta^1 = -\beta^2. \tag{105}$$

3.5. Dam break test cases

In this section we present comparisons among exact solutions of the Riemann problem for system (40), neglecting friction, and numerical solutions. The exact solution corresponding to a frictionless dambreak flow over a discontinuous bed is detailed in Appendix C. The results are presented in the form of plots of the total depth, mean discharge, Froude number and energy per unit weight or head. The examples are chosen to represent different combinations of wave patterns. The acceleration due to gravity is set equal to $g = 9.8 \text{ m}^2/\text{s}$. In all cases the bottom step is positioned at $x = 0$ and has a variable height. In all cases $\Delta x = 1$ and $\text{CFL} = 1$. When applying the reconstruction technique of the weak solution proposed in (104) and (105) no difference in the solution has been observed when comparing with the original solution.

Test case 1 is a dam-break type problem, with a combination of rarefaction and shock waves. The initial condition consists of two columns of water of different height and zero velocity. The solution, presented in Fig. 11, contains a left moving rarefaction wave, a stationary shock at the step and a right-moving shock wave. The presence of the step leads to a reduction of the total water height running to the right as compared to the flat bottom case. This reduction is due to the stationary shock, which dissipates part of the energy of the shock wave. All three approximations of the pressure term, $(p_b/\rho_w)^a$, $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$, provide results of similar accuracy, for the total depth, mean velocity, Froude number and energy. The options $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$ in particular overlap completely so that they cannot be distinguished in the plot. The position of

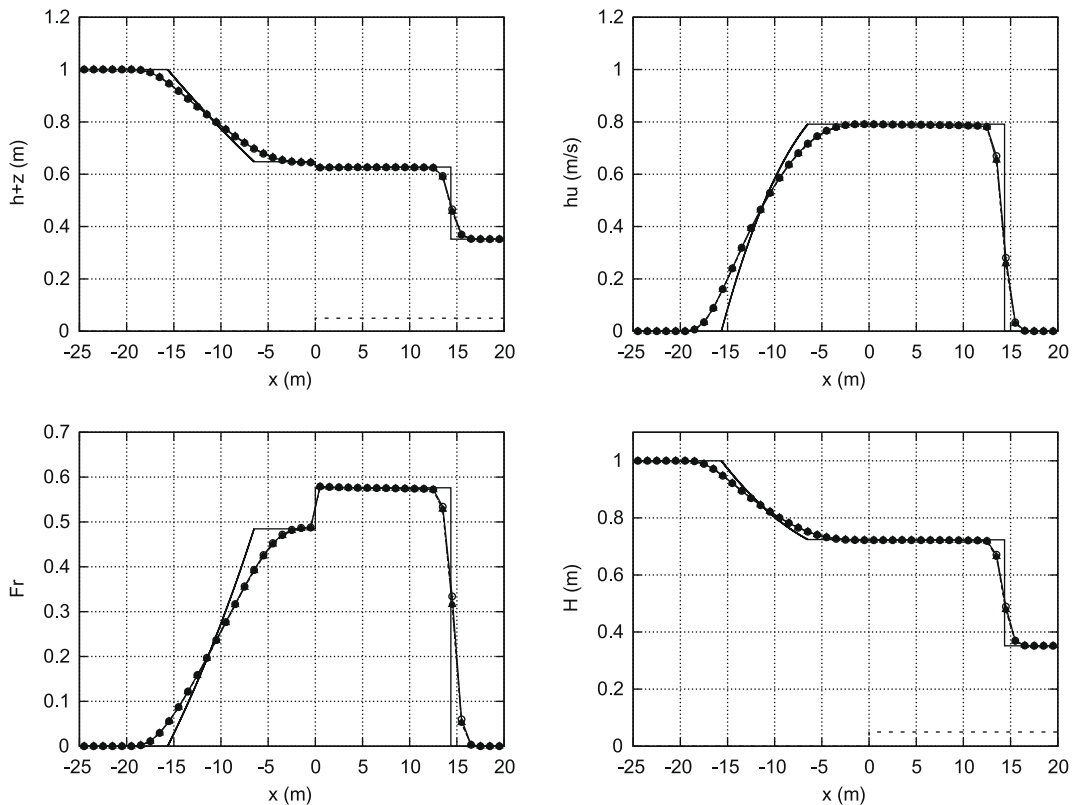


Fig. 11. Test case 1: comparison between exact (—) and numerical solutions at $t=5 \text{ s}$ obtained with approaches $(p_b/\rho_w)^a$ (—○—), $(p_b/\rho_w)^b$ (—●—) and $(p_b/\rho_w)^c$ (—△—).

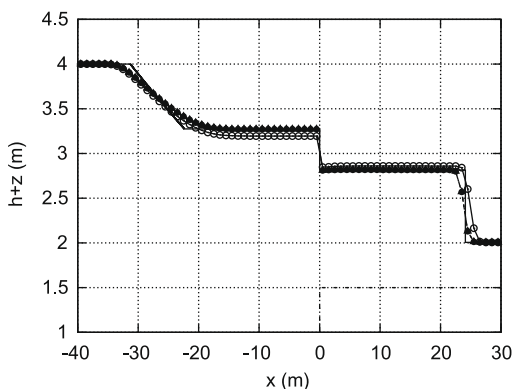
the fan expansion and the shock are correct and the discharge does not present oscillations in the origin. The correct behavior of the numerical scheme is explained attending to the characteristics of the weak solution, that in a subcritical case, provides a constant value of discharge for the two inner regions \mathbf{U}_i^* and \mathbf{U}_{i+1}^{**} , according to the exact solution (see Table 2).

Test case 2 is also a dam-break type problem, with a combination of rarefaction and shock waves. The initial condition consists of two columns of water of different heights and velocity on the left side. The solution, presented in Fig. 12, contains a left moving rarefaction wave, a stationary shock at the step and a right-moving shock wave. The numerical solution for approximation $(p_b/\rho_w)^a$ differs strongly from the analytical solution for all the plotted variables, and also provides an increment of the total water height. Approximations of the pressure term, $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$, provide the same results, leading to a correct description of the total depth, the mean discharge, Froude number and compute the energy dissipation correctly.

Test case 3 is a two shock case with a convergent flow. The solution is presented in Fig. 13, and contains a left-moving shock, a stationary shock at the step and a right-moving shock wave. As in the previous example, the step acts as an energy dissipation mechanism. All three approximations of the pressure term, $(p_b/\rho_w)^a$, $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$, provide results of similar accuracy, for the total depth, mean discharge and Froude number, as this case corresponds to a relatively mild slope.

Table 2
Summary of test cases.

Test case	h_L	h_R	u_L	u_R	z_L	z_R
1	1.0	0.30179953	0.0	0.0	0.0	0.05
2	4.0	0.50537954	0.1	0.0	0.0	1.5
3	2.5	2.49977381	1.5	0.0	0.0	0.25
4	1.5	0.16664757	2.0	0.0	0.0	2.0
5	1.0	0.04112267	0.2	0.0	0.25	0.0
6	0.6	0.02599708	0.35	0.0	1.2	0.0
7	1.1	0.49457729	4.9	6.50	0.2	0.0
8	1.5	0.0	2.2862	0.0	0.0	4.0
9	1.5	0.0	4.5	0.0	0.0	4.0
10	1.5	0.0	-2.5	0.0	0.0	1.0
11	1.5	0.0	-5.0	0.0	0.0	1.0



)

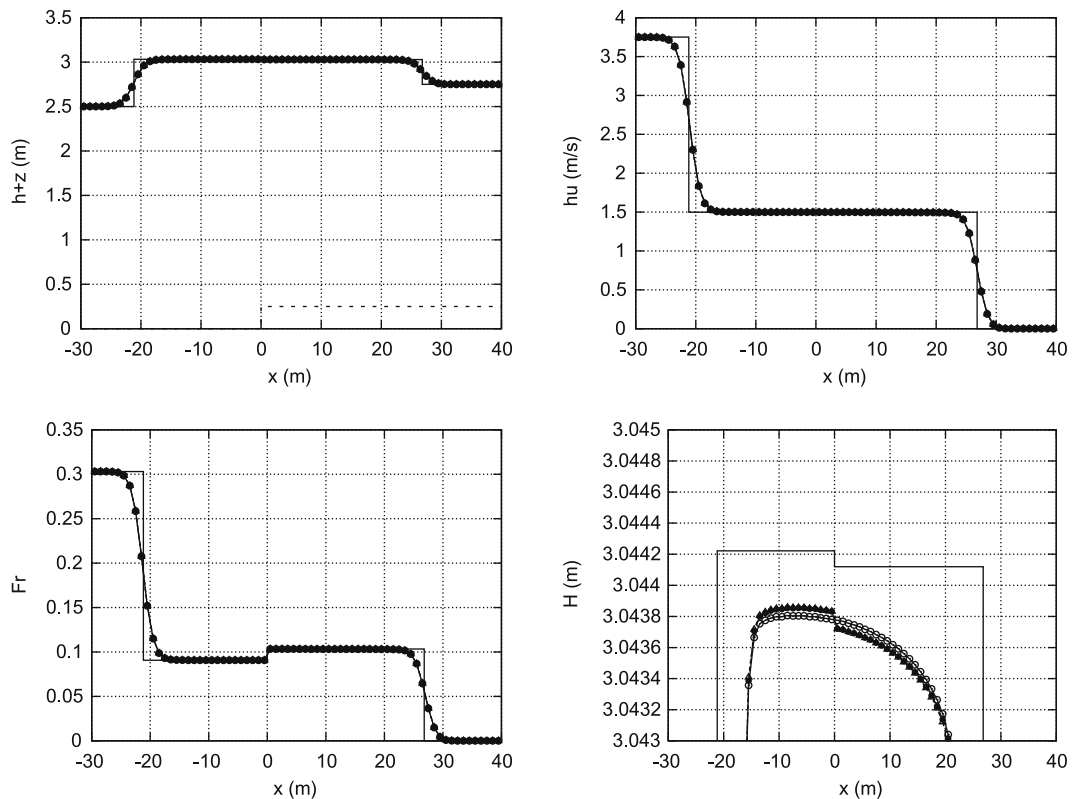


Fig. 13. Test case 3: comparison between exact (—) and numerical solutions at $t=5$ s obtained with approaches $(p_b/\rho_w)^a$ (—○—), $(p_b/\rho_w)^b$ (—●—) and $(p_b/\rho_w)^c$ (—△—).

The strongest differences appear in the total water height, with $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$ providing the most energy dissipative solutions.

Test case 4 is also a two shock case with a convergent flow, with an initial discontinuity in the water depth. The performance of the numerical scheme in cases with this type of discontinuity is of major importance in practice. As in test case 3, the solution, depicted in Fig. 14, contains a left-moving shock, a stationary shock at the step and a right-moving shock wave. All solutions provide an adequate description of the water discharge at the dam break position, $x=0$ m, but the differences among the results for the rest of variables are noticeable. Approach $(p_b/\rho_w)^a$ lead to unphysical result for the total water height, while approach $(p_b/\rho_w)^b$, strongly overestimates the velocity of the right moving shock and underestimates the velocity of the left moving shock. The hybrid option $(p_b/\rho_w)^c$ combines the best properties of both approaches leading to superior results, not only providing an energy dissipating solution but also reaching better results for the rest of variables. The differences with the exact solution are attributable to the linearization of the source term when constructing the weak solution.

Test case 5 is a dam-break type problem, with a combination of rarefaction and shock waves, but in contrast with test cases 1 and 2, the initial condition consists of two columns of water of different heights, velocity in the left side and a backward step of height similar to the right column of water. The solution, presented in Fig. 15, contains a left moving rarefaction wave, a stationary shock at the step and a right-moving shock wave. The presence of the step dissipates part of the energy. Approximations of the pressure term $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$, that provide identical results, lead to a description of the total depth, the mean discharge, Froude number and compute the energy dissipation closer to the exact solution. Again, the differences still shown by approaches $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$ with respect to the exact solution are attributable to the linearization of the source term when constructing the weak solution. Test case 6 is similar to test case 5, but with a stronger discontinuity in the water elevation, and a very thin layer of water on the right side. The solution is presented in Fig. 16. Only approximations of the pressure term $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$, that provide identical results, lead to a correct description of the total depth, the mean discharge, Froude number and compute the energy dissipation. Approach $(p_b/\rho_w)^a$, provides a solution totally inadequate and distorted if compared with the analytical case.

In test case 7 a supercritical motion from left to right is considered. The presence of the step introduces no limitation in the signal propagation upstream, and its effect is in dissipating energy by the stationary shock at the step. The solution, presented in Fig. 17, contains a stationary shock at the step, a right moving rarefaction wave, and a right-moving shock wave. Only approximations of the pressure term $(p_b/\rho_w)^b$ and $(p_b/\rho_w)^c$, that provide identical results, lead to a correct description

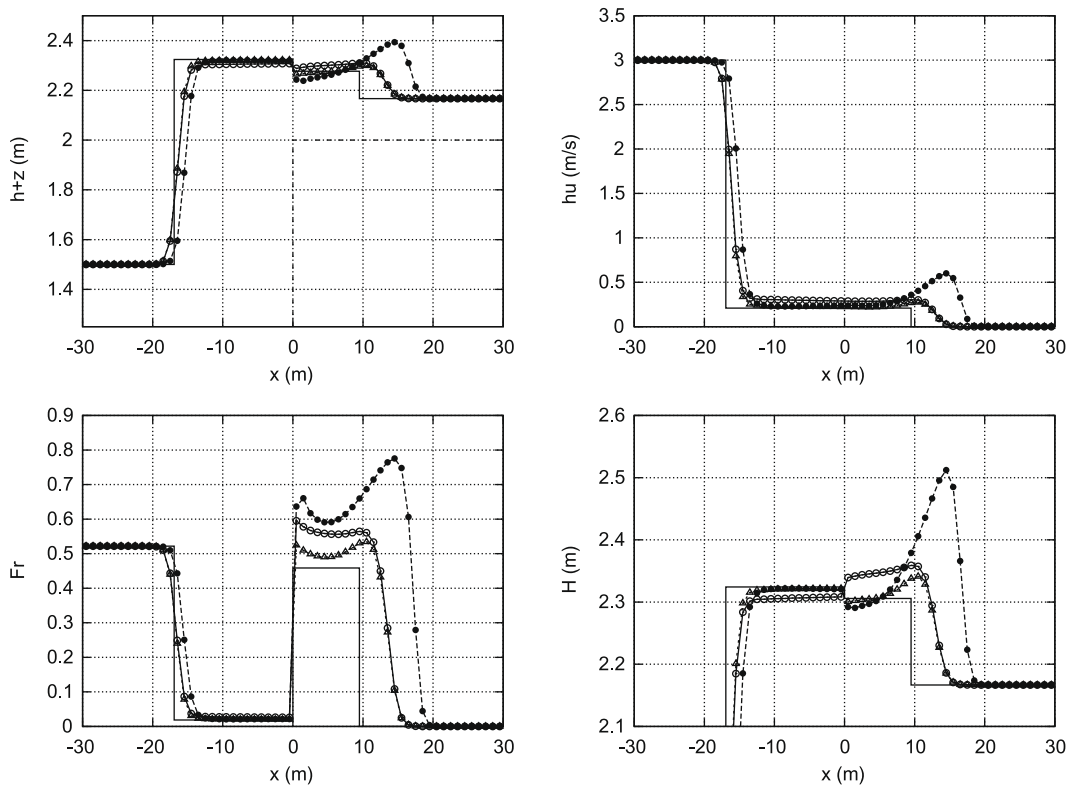


Fig. 14. Test case 4: comparison between exact (—) and numerical solutions at $t=5s$ obtained with approaches $(p_b/\rho_w)^a$ (—○—), $(p_b/\rho_w)^b$ (—●—) and $(p_b/\rho_w)^c$ (—△—).

of the total depth, the mean discharge, Froude number and compute the energy dissipation correctly. Approach $(p_b/\rho_w)^a$, provides an unphysical solution for the total water height. For all approaches the computed discharge is kept constant and equal to the left value, $(hu)_L$, until the solution reaches the rarefaction wave. This is explained attending to the definition of the weak solution for supercritical cases in (185), that provides an adequate value for the inner region, resulting in $(hu)_i^n = (hu)_{i+1}^*$ according to the exact solution.

In test cases 8–11 the performance of the numerical scheme of RPs characterized by a fixed in time wet/dry position are analysed. Depending on the approach of the source term the flow develops in both regions of the plain. This is avoided if the weak solution provides null or negative values of water depth in the initially dry side of the RP. In this context, the performance of the energy dissipative approach $(p_b/\rho_w)^{c,w}$ is of interest.

In test case 8 a subcritical flow encounters a wall and is reflected. The solution, presented in Fig. 18, contains a left-moving shock that links the right moving water with a region of still water ending at the stationary shock at $x=0$. The bed elevation on the right side is greater than the maximum allowable water depth in the wet region at rest. Only approximations of the pressure term $(p_b/\rho_w)^a$, $(p_b/\rho_w)^c$ and $(p_b/\rho_w)^{c,w}$ provide a correct solution in the left side, while $(p_b/\rho_w)^b$, develops in both sides leading to incorrect solutions. Option $(p_b/\rho_w)^{c,w}$ provides more accurate results for both water depth and flow discharge. The same results are observed for test case 9, Fig. 19, where a supercritical flow encounters a wall and is reflected.

In test case 10 a subcritical flow moving to the left generates a left moving rarefaction connecting the left state with a state of water at rest, with a water depth smaller than the depth of the wall defined in the other side. The solution is presented in Figs. 20 and 21. Approximations of the pressure term $(p_b/\rho_w)^a$ and $(p_b/\rho_w)^c$ develop correctly in the left side of the (x,t) plane, although present an oscillatory behavior in the zero velocity region. The results for approach $(p_b/\rho_w)^{c,w}$ do not present any oscillation in the region of zero velocity but estimate less accurately the velocities at the tail and the head of the rarefaction. Again $(p_b/\rho_w)^b$ provides an incorrect solution.

The situation is much more interesting in test case 11, where supercritical flow conditions are enforced in the left side. Again $(p_b/\rho_w)^b$ provides an incorrect solution and $(p_b/\rho_w)^a$, $(p_b/\rho_w)^c$ lead to similar results. Results given by approach $(p_b/\rho_w)^{c,w}$ follow the same tendency as in test case 10.

According to the hypothesis assumed to derive the exact solution, the solutions in test cases 8–11 should be independent of the actual height of the dry step z_R . That is, if the value of the dry bed level z_R decreases until a value equal to the water depth in the region of zero velocity, that will be referred to as $h_{L,0}$, the solution remains equal. However, due to the approximation used to define the bed discontinuity, this has an influence, and before z_R reaches the limit $h_{L,0}$, numerical solutions

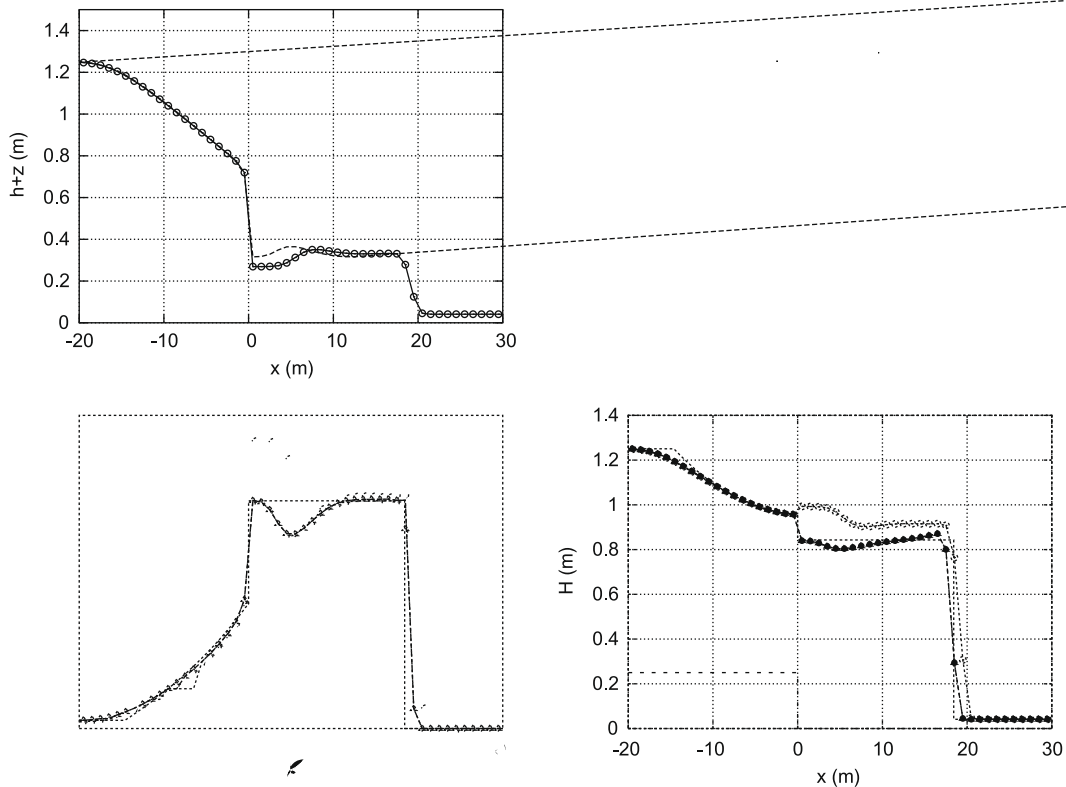


Fig. 15. Test case 5: comparison between exact (—) and numerical solutions at $t=5$ s obtained with approaches $(p_b/\rho_w)^a$ (—○—), $(p_b/\rho_w)^b$ (—●—) and $(p_b/\rho_w)^c$ (—△—).

evolve in both regions. Table 3 shows the numerical limits that can be reached by the value of z_R before the solution is altered. For test cases 8 and 9, among the approaches described, $(p_b/\rho_w)^{c,w}$ provides the most accurate limiting value, denoted $z_R^{c,w}$, while approaches $(p_b/\rho_w)^a$ and $(p_b/\rho_w)^c$ require larger values to ensure a correct solution. If $(p_b/\rho_w)^b$ is used, the solution fails for all values of z_R^b . For test case 10 in Table 3 shows that $(p_b/\rho_w)^{c,w}$ leads to the best results when decreasing the value for z_R and that $(p_b/\rho_w)^b$ fails in all cases. It is noticeable that, in case 11, approaches $(p_b/\rho_w)^a$, $(p_b/\rho_w)^c$ and $(p_b/\rho_w)^{c,w}$ lead to similar limits for z_R .

3.6. Dam break over an adverse slope

In the following test case the numerical solution of a dam break case, over an adverse dry bed is analyzed. A frictionless, confined domain 250 m long, presents the following initial conditions and bathymetry:

$$\begin{aligned}
 h(x, 0) &= \begin{cases} 0.8 & \text{if } x < 50, \\ 0 & \text{if } x > 50, \end{cases} \quad u(x, 0) = u_o(x) = 0, \\
 z(x) &= \begin{cases} 1.2 & \text{if } x < 50, \\ 0.01(x - 50) & \text{if } x > 50. \end{cases}
 \end{aligned} \tag{106}$$

Although no exact solution is presented for this test case it is useful to note the effect of the reconstruction techniques in (104) and (105). In this test case the left column of water must fall, leading to a left rarefaction in the flat bed region and to a flooding wave over an adverse initially dry slope. In all cases $\Delta x = 1$ m, CFL = 1 approach $(p_b/\rho_w)^c$ is used.

Fig. 22(a) and (b) show the bed level surface and water discharge at time $t = 45$ s given the numerical scheme (78), when computed using the original weak solutions (182)–(188) (—○—) and computed using modified weak solutions (—△—) generated through the reconstruction techniques in (104) and (105). The results for both weak solutions present no differences for the rarefaction wave and the advancing front. Fig. 24 displays time step evolution in time with the original weak solution (—○—) and with the reconstructed weak solution (—△—). It can be seen that the time step for both solutions is equal during approximately the first 45 seconds, as the source terms do not play a major role in the stability region, governed by Δt^2 .

As time evolves the water level in the flat left region decreases. Fig. 23(c) shows that at time $t = 50$ s the original weak solution provides smaller values of water depth than the reconstructed weak solution in the nodes located after and before

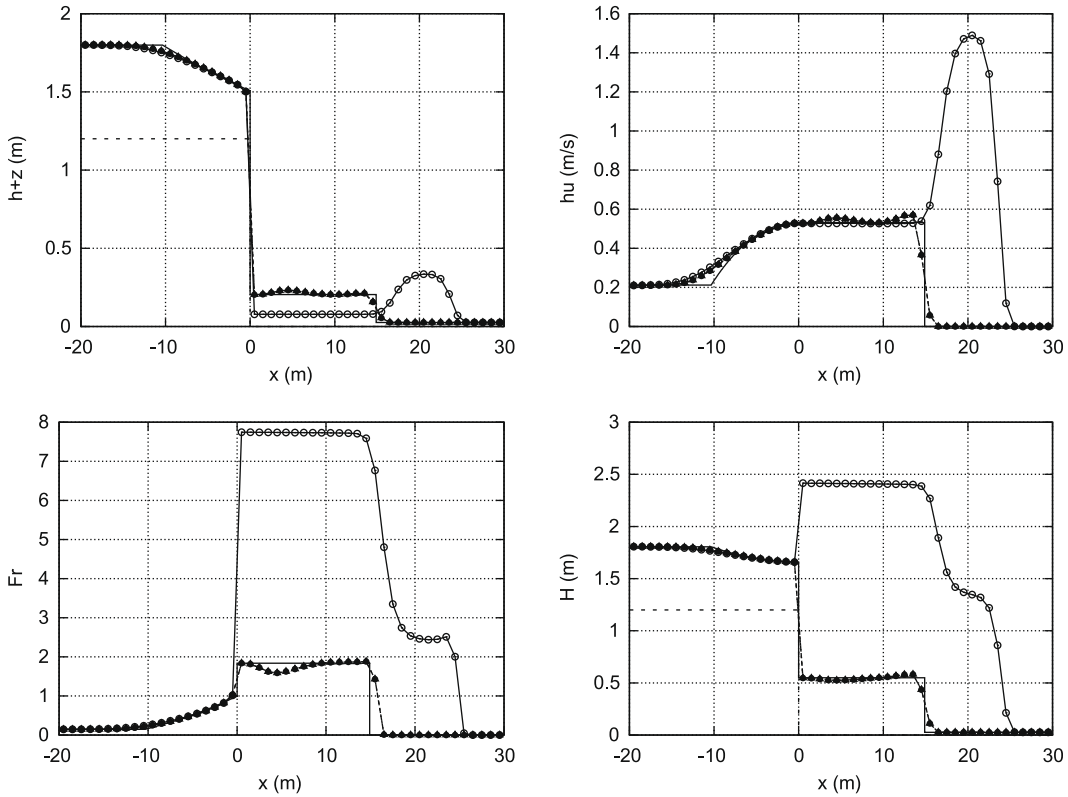


Fig. 16. Test case 6: comparison between exact (–) and numerical solutions at $t = 5$ s obtained with approaches $(p_b/\rho_w)^a$ (–○–), $(p_b/\rho_w)^b$ (–●–) and $(p_b/\rho_w)^c$ (–△–).

the ending of the flat bed, and how this difference is more noticeable at time $t = 55$ s, as depicted in Fig. 23(e). Also, less oscillatory results for the water discharge are computed in the same location, as seen in Fig. 23(d) for $t = 50$ s and (f) $t = 55$ s. Almost unnoticeable differences appear in the advancing front.

Fig. 24 shows that, when both approximate solutions begin to diverge, the time step size also does so. For times approximately between $t = 45$ and $t = 50$ s, smaller values of time step are required for the original weak solution if compared with the ones required by the modified solution. For longer times, the time step for the original weak solution becomes several orders of magnitude smaller than the time step for the reconstructed solution, as the source terms influence the stability region, rendering the simulation of longer times than $t = 55$ impossible. On the other hand, the reconstructed solution allows a time step size of constant order of magnitude.

In the previous section approach $(p_b/\rho_w)^{c,w}$ lead to accurate results for reflection waves and rarefactions in stationary wet/dry problems. Fig. 23 shows the results given by the energy dissipative approach $(p_b/\rho_w)^{c,w}$ compared with those given by approach $(p_b/\rho_w)^c$. As expected, the main difference appears in the velocity of the advancing front.

4. 2D systems of equations with source terms

The previous ideas are next extended to 2D hyperbolic nonlinear systems of equations with source terms of the form

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} + \frac{\partial \mathbf{G}(\mathbf{U})}{\partial y} = \mathbf{S}(\mathbf{U}, \mathbf{x}, \mathbf{y}). \tag{107}$$

System (107) is time dependent, nonlinear, and contains source terms. Under the hypothesis of dominant advection it can be classified and numerically dealt with as belonging to the family of hyperbolic systems. The mathematical properties of (107) include the existence of a Jacobian matrix, \mathbf{J}_n , of the flux normal to a direction given by the unit vector \mathbf{n} , $\mathbf{E}\mathbf{n}$, with $\mathbf{E} = \mathbf{F}\mathbf{n}_x + \mathbf{G}\mathbf{n}_y$, defined as

$$\mathbf{J}_n = \frac{\partial \mathbf{E}\mathbf{n}}{\partial \mathbf{U}} = \frac{\partial \mathbf{F}}{\partial \mathbf{U}} n_x + \frac{\partial \mathbf{G}}{\partial \mathbf{U}} n_y. \tag{108}$$

This Jacobian can be used to form the basis of the upwind numerical discretization that will be outlined in next section.

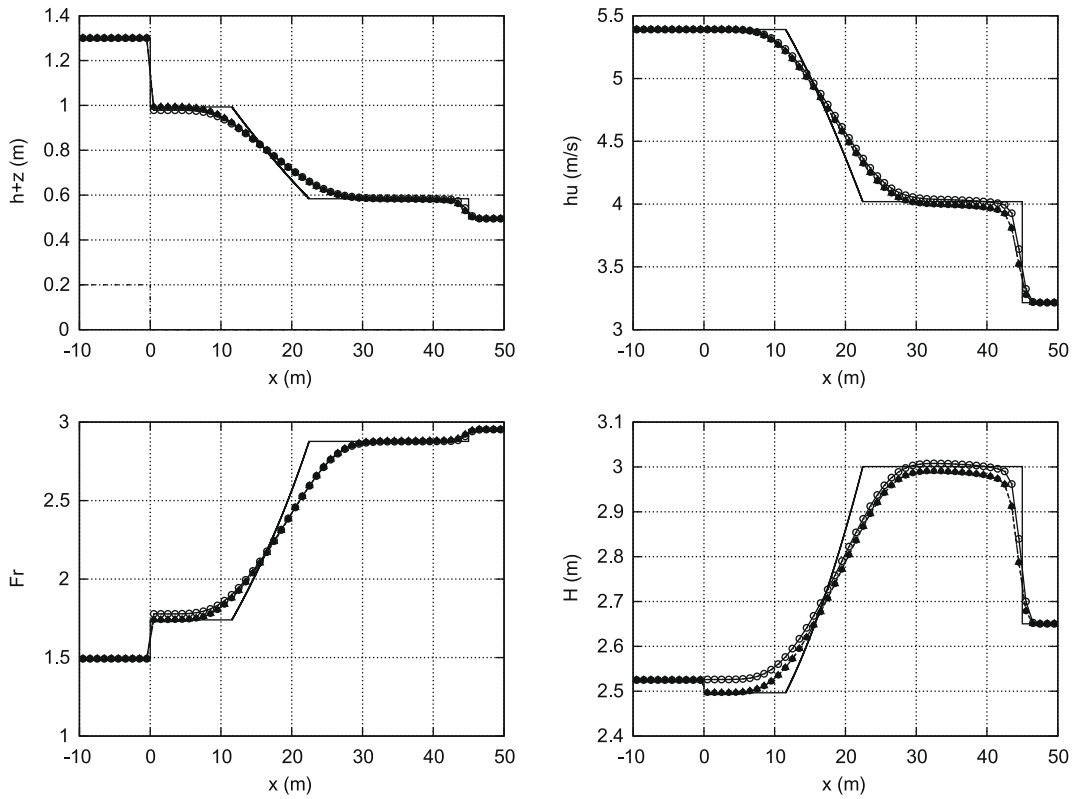


Fig. 17. Test case 7: comparison between exact (–) and numerical solutions at $t=5$ s obtained with approaches $(p_b/\rho_w)^a$ (–○–), $(p_b/\rho_w)^b$ (–●–) and $(p_b/\rho_w)^c$ (–△–).

4.1. Approximate Riemann solution

To introduce the finite volume scheme, (107) is integrated in a volume or grid cell Ω :

$$\frac{\partial}{\partial t} \int_{\Omega} \mathbf{U} d\Omega + \int_{\Omega} (\nabla \cdot \mathbf{E}) d\Omega = \int_{\Omega} \mathbf{S} d\Omega \tag{109}$$

and applying the Gauss theorem becomes

$$\frac{\partial}{\partial t} \int_{\Omega} \mathbf{U} d\Omega + \oint_{\partial\Omega} \mathbf{E} \mathbf{n} dl = \int_{\Omega} \mathbf{S} d\Omega, \tag{110}$$

where $\mathbf{n} = (n_x, n_y)$ is the outward unit normal vector to the volume Ω .

In order to obtain a numerical solution of system (107) we divide the domain in computational cells, Ω_i , using a mesh fixed in time. Let Δt be the time step and $t^n = n\Delta t$ a generic time. Assuming the usual notation we indicate with \mathbf{U}_i^n the cell-average value of the solution $\mathbf{U}(x, y, t)$ for the i th cell at time t^n :

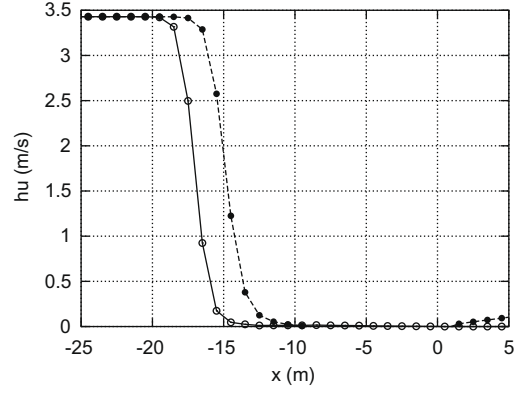
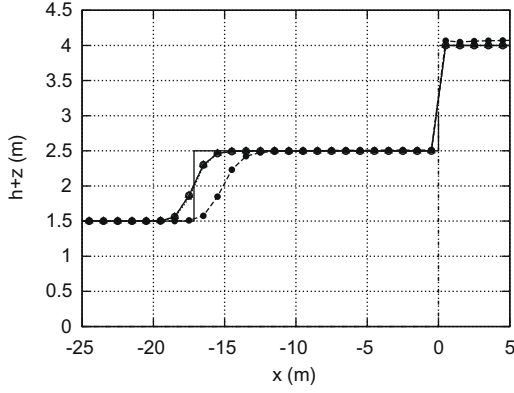
$$\mathbf{U}_i^n = \frac{1}{A_i} \int_{\Omega_i} \mathbf{U}(x, y, t^n) dx, \tag{111}$$

where A_i is the cell area. Assuming a piecewise representation of the conserved variables (110) is written as

$$\frac{\partial}{\partial t} \int_{\Omega_i} \mathbf{U} d\Omega + \sum_{k=1}^{NE} \mathbf{E}_j \mathbf{n}_k l_k = \int_{\Omega_i} \mathbf{S} d\Omega, \tag{112}$$

with \mathbf{E}_j the value of the function \mathbf{E} at the neighbouring cell j connected through the edge k , $\mathbf{n}_k = (n_x, n_y)$, is the outward unit normal vector to the cell edge k , l_k is the corresponding edge length and NE is the number of edges in cell i . Considering that the quantity \mathbf{E}_i is uniform per cell i and that

$$\sum_{k=1}^{NE} \mathbf{n}_k l_k = \mathbf{0}. \tag{113}$$



Eq. (112) is written as

$$\frac{\partial}{\partial t} \int_{\Omega_i} \mathbf{U} d\Omega + \sum_{k=1}^{NE} (\delta \mathbf{E})_k \mathbf{n}_k l_k = \int_{\Omega_i} \mathbf{S} d\Omega, \quad (114)$$

with $\delta \mathbf{E} = \mathbf{E}_j - \mathbf{E}_i$.

In the Roe approach, the solution of each RP is obtained from the exact solution of a locally linearized problem. In the 2D framework the solution is obtained reducing each RP at each k edge to a 1D Riemann problem projected onto the direction \mathbf{n} . The linearized solution must fulfill the Consistency Condition. In the 2D case the integral of the approximate solution $\hat{\mathbf{U}}(x', t)$ of the k linearized RP over a suitable control volume must be equal to the integral of the exact solution $\mathbf{U}(x', t)$ over the same control volume, with x' the coordinate normal to the cell edge k , Fig. 25. Then in each k Riemann problem with initial values $\mathbf{U}_i, \mathbf{U}_j$, in a time interval $[0, 1]$ and a space interval $[-X', X']$, where

$$-X' \leq \lambda_{\min}, \quad X' \geq \lambda_{\max} \quad (115)$$

and $\lambda_{\min}, \lambda_{\max}$ the positions of the slowest and the fastest wave at $t = 1$, in a k edge, the solution $\hat{\mathbf{U}}(x', 1)$ at time $t = 1$ must satisfy the following property:

$$\int_{-X'}^{+X'} \hat{\mathbf{U}}(x', 1) dx' = \int_{-X'}^{+X'} \mathbf{U}(x', 1) dx', \quad (116)$$

so using (114) the Consistency Condition becomes:

$$\int_{-X'}^{+X'} \hat{\mathbf{U}}(x', 1) dx' = X'(\mathbf{U}_R + \mathbf{U}_L) - \delta \mathbf{E} \mathbf{n} + \int_0^1 \int_{-X'}^{+X'} \mathbf{S} dx' dt. \quad (117)$$

Since the source terms are not necessarily constant in time, we assume the following time linearization of the Consistency Condition:

$$\int_{-X'}^{+X'} \hat{\mathbf{U}}(x', 1) dx' = X(\mathbf{U}_R + \mathbf{U}_L) - (\delta \mathbf{E} - \mathbf{T})_k \mathbf{n}_k, \quad (118)$$

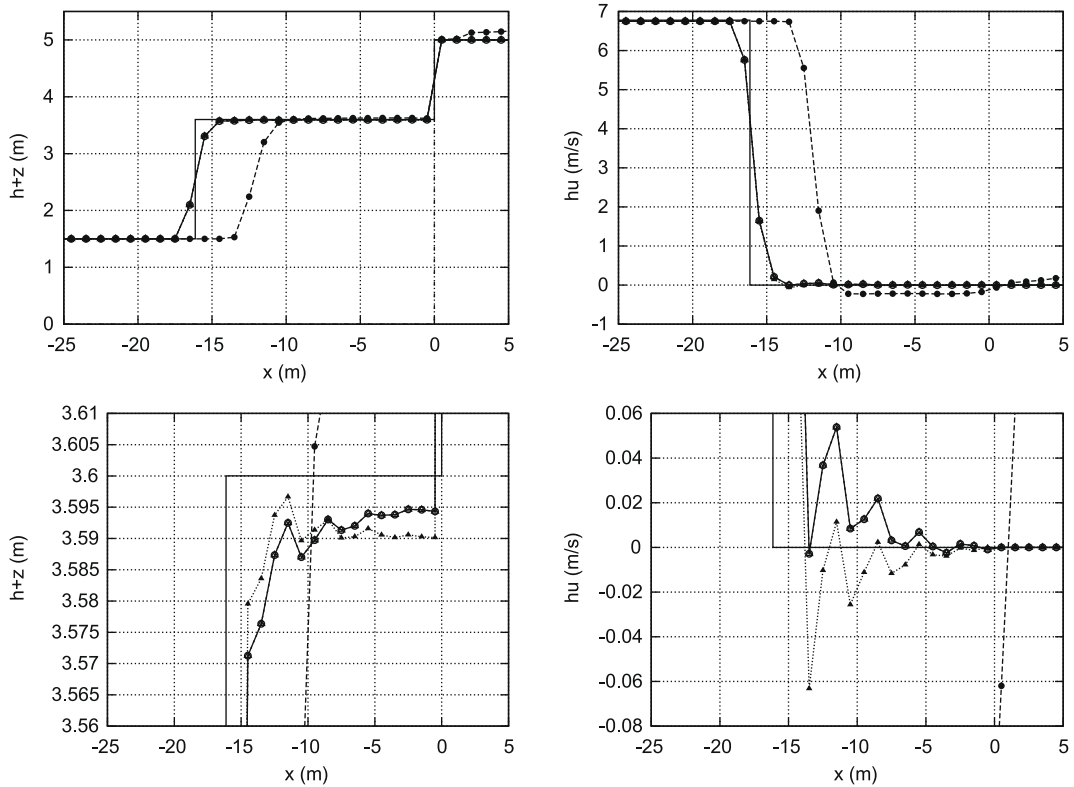


Fig. 19. Test case 9: comparison between exact (—) and numerical solutions at $t=5$ s obtained with approaches $(p_b/\rho_w)^a$ (---), $(p_b/\rho_w)^b$ (—●—), $(p_b/\rho_w)^c$ (---▲---) and $(p_b/\rho_w)^{c,w}$ (—▲—).

where following previous work, [20]

$$\int_{-X'}^{+X'} \mathbf{S}(x', 0) dx' = (\mathbf{Tn})_k^n, \tag{119}$$

where \mathbf{T} is a suitable numerical source matrix. This enables the following formulation of (114)

$$\frac{\partial}{\partial t} \int_{\Omega} \mathbf{U} d\Omega_i + \sum_{k=1}^{NE} (\delta \mathbf{E} - \mathbf{T})_k \mathbf{n}_k l_k = 0, \tag{120}$$

that is approximated by using the following linear problem

$$\begin{aligned} \frac{\partial}{\partial t} \int_{\Omega} \widehat{\mathbf{U}} d\Omega_i + \sum_{k=1}^{NE} \mathbf{J}_{\mathbf{n},k}^* \delta \widehat{\mathbf{U}}_k l_k &= 0, \\ \widehat{\mathbf{U}}(x', 0)_k &= \begin{cases} \mathbf{U}_i & \text{if } x' < 0, \\ \mathbf{U}_j & \text{if } x' > 0. \end{cases} \end{aligned} \tag{121}$$

Integrating the previous PDE system over the same control volume the following expression is obtained in each k edge

$$\int_{-X'}^{+X'} \widehat{\mathbf{U}}(x', 1) dx' = X(\mathbf{U}_i + \mathbf{U}_j) - \mathbf{J}^*(\mathbf{U}_j - \mathbf{U}_i) \tag{122}$$

and since we want to satisfy (118), the constraint that follows is:

$$(\delta \mathbf{E} - \mathbf{T})_k \mathbf{n}_k = \widetilde{\mathbf{J}}^*(\mathbf{U}_{i+1} - \mathbf{U}_i). \tag{123}$$

Due to the non-linear character of the flux \mathbf{E} , the definition of an approximated Jacobian matrix, $\widetilde{\mathbf{J}}_{\mathbf{n},k}$, allows for a local linearization

$$\delta(\mathbf{E}\mathbf{n})_k = \widetilde{\mathbf{J}}_{\mathbf{n},k} \delta \mathbf{U}_k \tag{124}$$

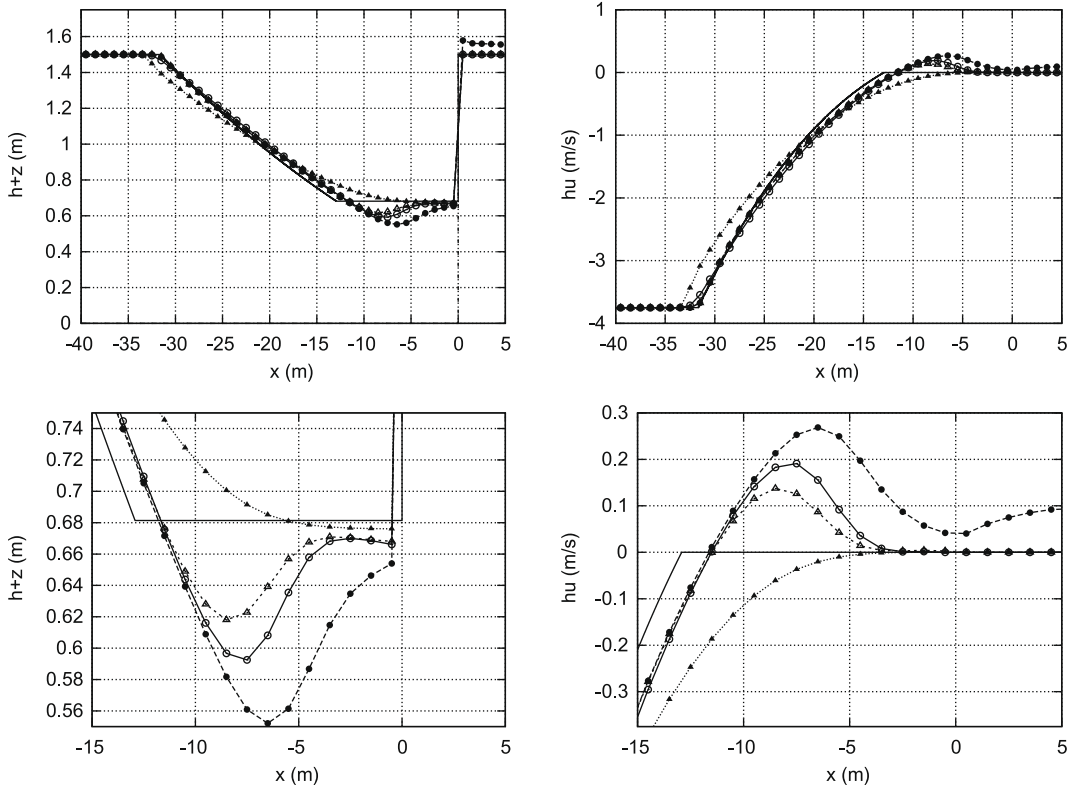


Fig. 20. Test case 10: comparison between exact (—) and numerical solutions at $t=5$ s obtained with approaches $(p_b/\rho_w)^a$ (—○—), $(p_b/\rho_w)^b$ (—●—), $(p_b/\rho_w)^c$ (—△—) and $(p_b/\rho_w)^{c,w}$ (—▲—).

and is exploited here [14]. This approach provides a set of three real eigenvalues $\tilde{\lambda}_k^m$ and eigenvectors $\tilde{\mathbf{e}}_k^m$. Then, it is possible to define two approximate matrices $\tilde{\mathbf{P}} = (\tilde{\mathbf{e}}^1, \tilde{\mathbf{e}}^2, \tilde{\mathbf{e}}^3)$ and $\tilde{\mathbf{P}}^{-1}$ with the following properties

$$\tilde{\mathbf{J}}_{n,k} = \tilde{\mathbf{P}}_k \tilde{\Lambda}_k \tilde{\mathbf{P}}_k^{-1}. \tag{125}$$

The difference in vector \mathbf{U} across the grid edge is projected onto the matrix eigenvectors basis and the same for the source term:

$$\delta \mathbf{U}_k = \tilde{\mathbf{P}}_k \mathbf{A}_k, \quad (\mathbf{Tn})_k = \tilde{\mathbf{P}}_k \mathbf{B}_k, \tag{126}$$

with $\mathbf{A}_k = (\alpha^1 \ \alpha^2 \ \alpha^3)_k^T$ and $\mathbf{B}_k = (\beta^1 \ \beta^2 \ \beta^3)_k^T$. Expressing all terms more compactly:

$$\delta(\mathbf{E}\mathbf{n})_k - (\mathbf{Tn})_k = \sum_{m=1}^{N_s} (\tilde{\lambda}^m \theta \alpha \tilde{\mathbf{e}})_k^m, \tag{127}$$

with

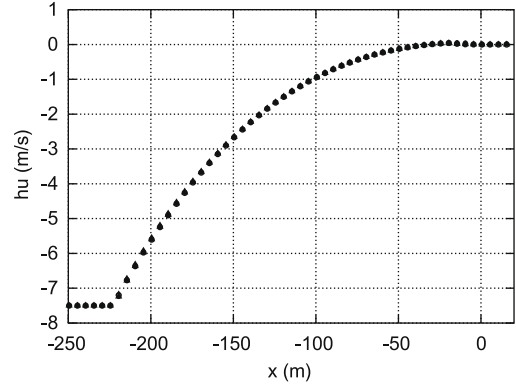
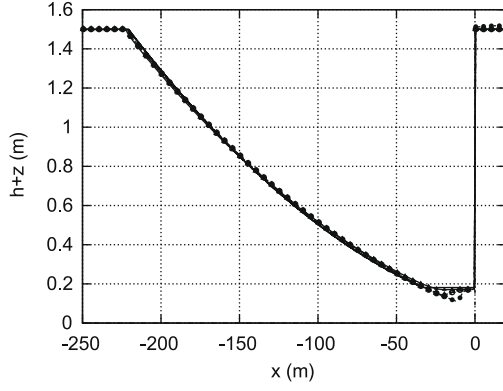
$$\theta_k^m = \left(1 - \frac{\beta}{\tilde{\lambda} \alpha}\right)_k^m \tag{128}$$

it is possible to define the desired matrix in (123)

$$\tilde{\mathbf{J}}_k^* = (\tilde{\mathbf{P}} \tilde{\Lambda}^* \tilde{\mathbf{P}}^{-1})_k, \tag{129}$$

with $\tilde{\Lambda}^* = \tilde{\Lambda} \Theta$, where $\tilde{\Lambda}_k$ is a diagonal matrix with eigenvalues $\tilde{\lambda}_k^{m,*}$ in the main diagonal and Θ_k is a diagonal matrix with θ_k^m in the main diagonal:

$$\tilde{\Lambda}_k = \begin{pmatrix} \tilde{\lambda}^1 & 0 & 0 \\ 0 & \tilde{\lambda}^2 & 0 \\ 0 & 0 & \tilde{\lambda}^3 \end{pmatrix}_k, \quad \Theta_k = \begin{pmatrix} \theta^1 & 0 & 0 \\ 0 & \theta^2 & 0 \\ 0 & 0 & \theta^3 \end{pmatrix}_k. \tag{130}$$



4.2. Application to the 2D shallow water equations

The two-dimensional shallow water equations, which represent mass and momentum conservation in a plane, can be obtained by depth-averaging the Navier–Stokes equations. Neglecting diffusion of momentum due to viscosity and turbulence, wind effects and the Coriolis term, they form a system of equations [1] as in (107), where

$$\mathbf{U} = (h, q_x, q_y)^T \quad (131)$$

are the conserved variables with h representing the water depth, $q_x = hu$ and $q_y = hv$, with (u, v) the depth averaged components of the velocity vector \mathbf{u} along the (x, y) coordinates respectively. The fluxes of these variables are given by:

$$\mathbf{F} = \left(q_x, \frac{q_y^2}{h} + \frac{1}{2}gh^2, \frac{q_x q_y}{h} \right)^T, \quad \mathbf{G} = \left(q_y, q_x q_y/h, \frac{q_x^2}{h} + \frac{1}{2}gh^2 \right)^T, \quad (132)$$

where g is the acceleration of the gravity. The source terms of the system are the bed slope and the friction terms:

$$\mathbf{S} = \left(0, \frac{p_{b,x}}{\rho_w} - \frac{\tau_{b,x}}{\rho_w}, \frac{p_{b,y}}{\rho_w} - \frac{\tau_{b,y}}{\rho_w} \right)^T, \quad (133)$$

where the bed slopes of the bottom level z in its differential form are

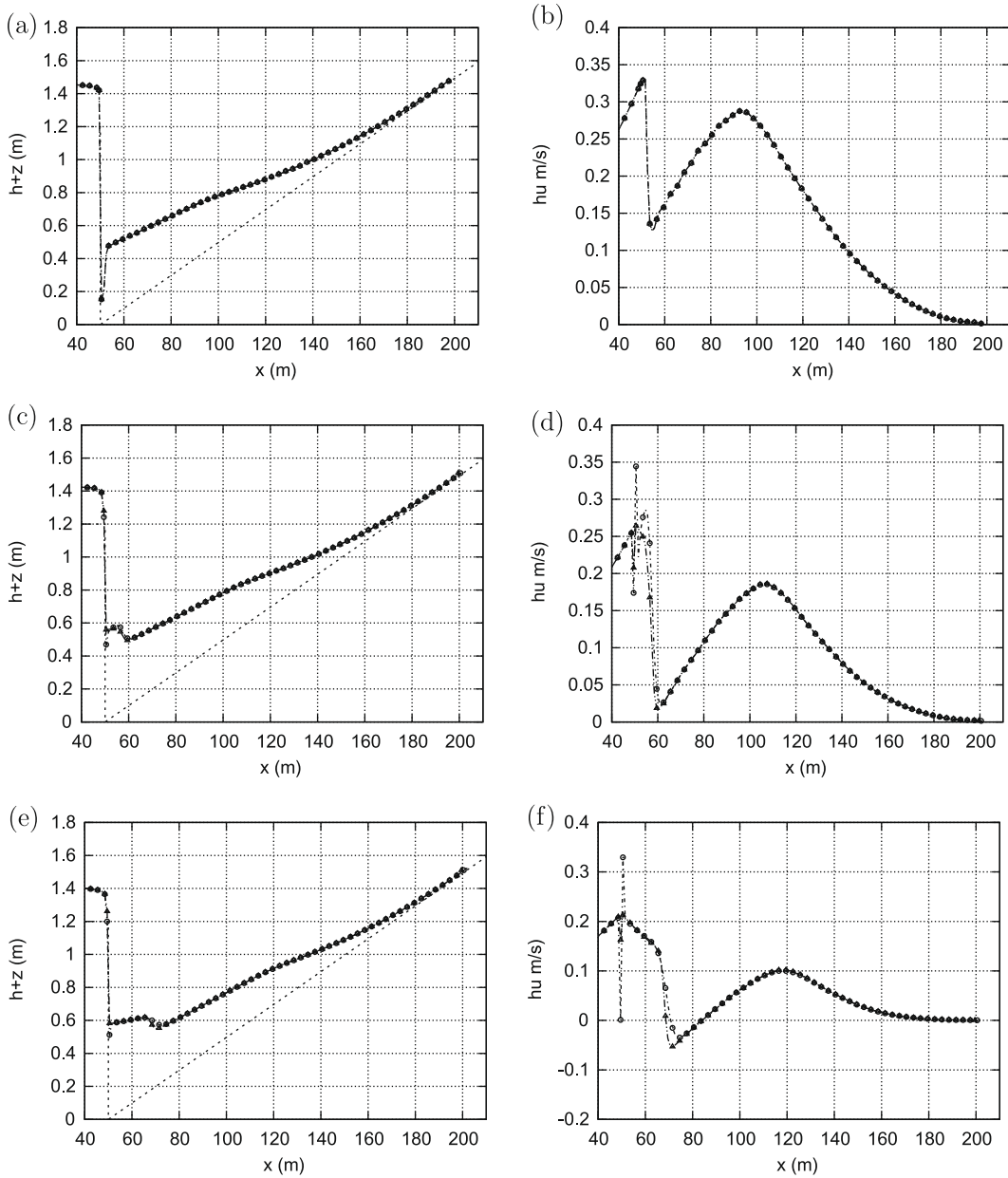


Fig. 22. Plot of the numerical water level surface and water discharge using approach $(p_b/\rho_w)^c$ (—○—) and $(p_b/\rho_w)^{c,w}$ (—△—) at times: (a) and (b) $t = 45$, (c) and (d) $t = 50$, (e) and (f) $t = 55$ s.

$$\frac{p_{b,x}}{\rho_w} = -gh \frac{\partial z}{\partial x}, \quad \frac{p_{b,y}}{\rho_w} = -gh \frac{\partial z}{\partial y} \tag{134}$$

and the friction losses are written in terms of the Manning's roughness coefficient n :

$$\frac{\tau_{b,x}}{\rho_w} = ghS_{fx}, \quad S_{fx} = \frac{n^2 u \sqrt{u^2 + v^2}}{h^{4/3}}, \quad \frac{\tau_{b,y}}{\rho_w} = ghS_{fy}, \quad S_{fy} = \frac{n^2 v \sqrt{u^2 + v^2}}{h^{4/3}}. \tag{135}$$

The approximate Jacobian $\tilde{J}_{n,k}$ for the homogeneous part is constructed with the following averaged variables [14]

$$\tilde{u}_k = \frac{u_i \sqrt{h_i} + u_j \sqrt{h_j}}{\sqrt{h_i} + \sqrt{h_j}}, \quad \tilde{v}_k = \frac{v_i \sqrt{h_i} + v_j \sqrt{h_j}}{\sqrt{h_i} + \sqrt{h_j}}, \quad \tilde{c}_k = \sqrt{g \frac{h_i + h_j}{2}} \tag{136}$$

leading to

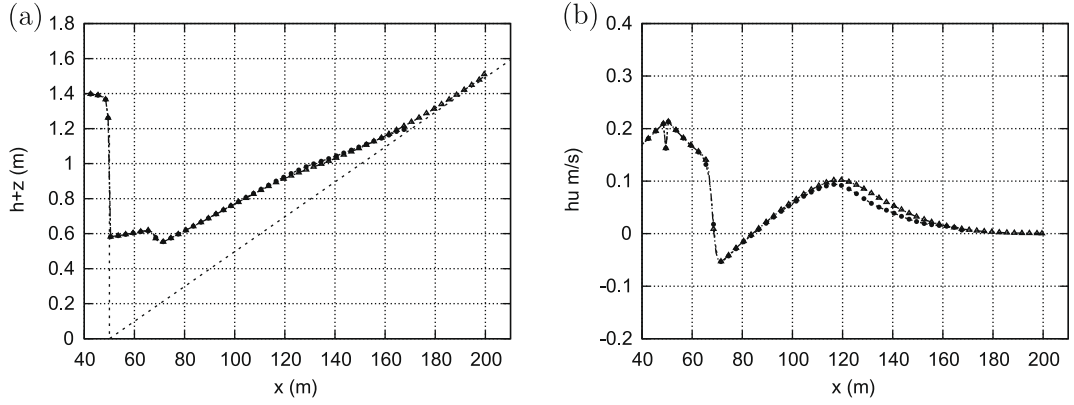
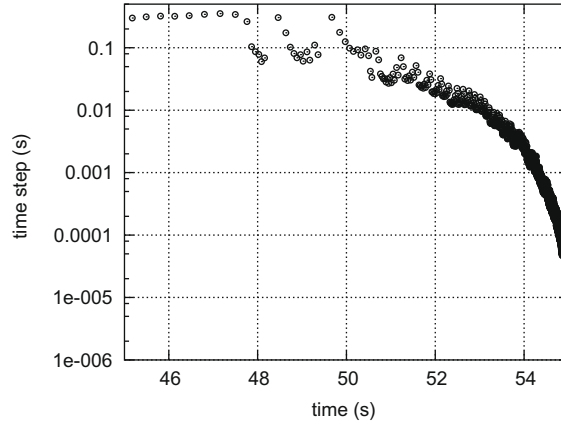


Fig. 23. Plot of the numerical water level surface and water discharge using the reconstructed weak solution with approach $(p_b/\rho_w)^c$ (— Δ —) and with approach $(p_b/\rho_w)^{c,w}$ (— \bullet —) at time $t = 55$ s.



$$\tilde{\lambda}_k^1 = (\tilde{\mathbf{u}}\mathbf{n} - \tilde{c})_k, \quad \tilde{\lambda}_k^2 = (\tilde{\mathbf{u}}\mathbf{n})_k, \quad \tilde{\lambda}_k^3 = (\tilde{\mathbf{u}}\mathbf{n} + \tilde{c})_k \quad (137)$$

and

$$\tilde{\mathbf{e}}_k^1 = \begin{pmatrix} 1 \\ \tilde{u} - \tilde{c}n_x \\ \tilde{v} - \tilde{c}n_y \end{pmatrix}_k, \quad \tilde{\mathbf{e}}_k^2 = \begin{pmatrix} 0 \\ -\tilde{c}n_y \\ \tilde{c}n_x \end{pmatrix}_k, \quad \tilde{\mathbf{e}}_k^3 = \begin{pmatrix} 1 \\ \tilde{u} + \tilde{c}n_x \\ \tilde{v} + \tilde{c}n_y \end{pmatrix}_k. \quad (138)$$

4.2.1. A four wave approximate Riemann solution

Four approximate solutions that satisfy (122) are proposed. These solutions for $\hat{\mathbf{U}}(\mathcal{X}', t)$ are governed by the celerities in $\tilde{\Lambda}_k$ and each one consists of five regions connected by four waves, one of them steady. The details are provided in [Appendix D](#).

Following Godunov’s method, the solutions of these RP’s are evolved for a time equal to the time step and the resulting solution is cell-averaged. The volume integral in the cell at time t^{n+1} is:

$$\mathbf{U}_i^{n+1} A_i = \mathbf{U}_i^n A_i - \sum_{k=1}^{NE} \sum_{m=1}^3 (\tilde{\lambda}^- \theta \alpha \tilde{\mathbf{e}})_k^m l_k \Delta t, \tag{139}$$

with $\tilde{\lambda}_k^{\pm, m} = \frac{1}{2} (\tilde{\lambda} \pm |\tilde{\lambda}|)_k^m$.

Again, algebraic manipulation converts (139) to an equivalent numerical flux-based finite volume scheme [8],

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_k \mathbf{E}_k^* \mathbf{n}_k l_k \frac{\Delta t}{A_i} + \sum_k (\mathbf{Tn})_k^- l_k \frac{\Delta t}{A_i} \tag{140}$$

in which the numerical flux (denoted by an asterisk) for first-order upwinding is given by

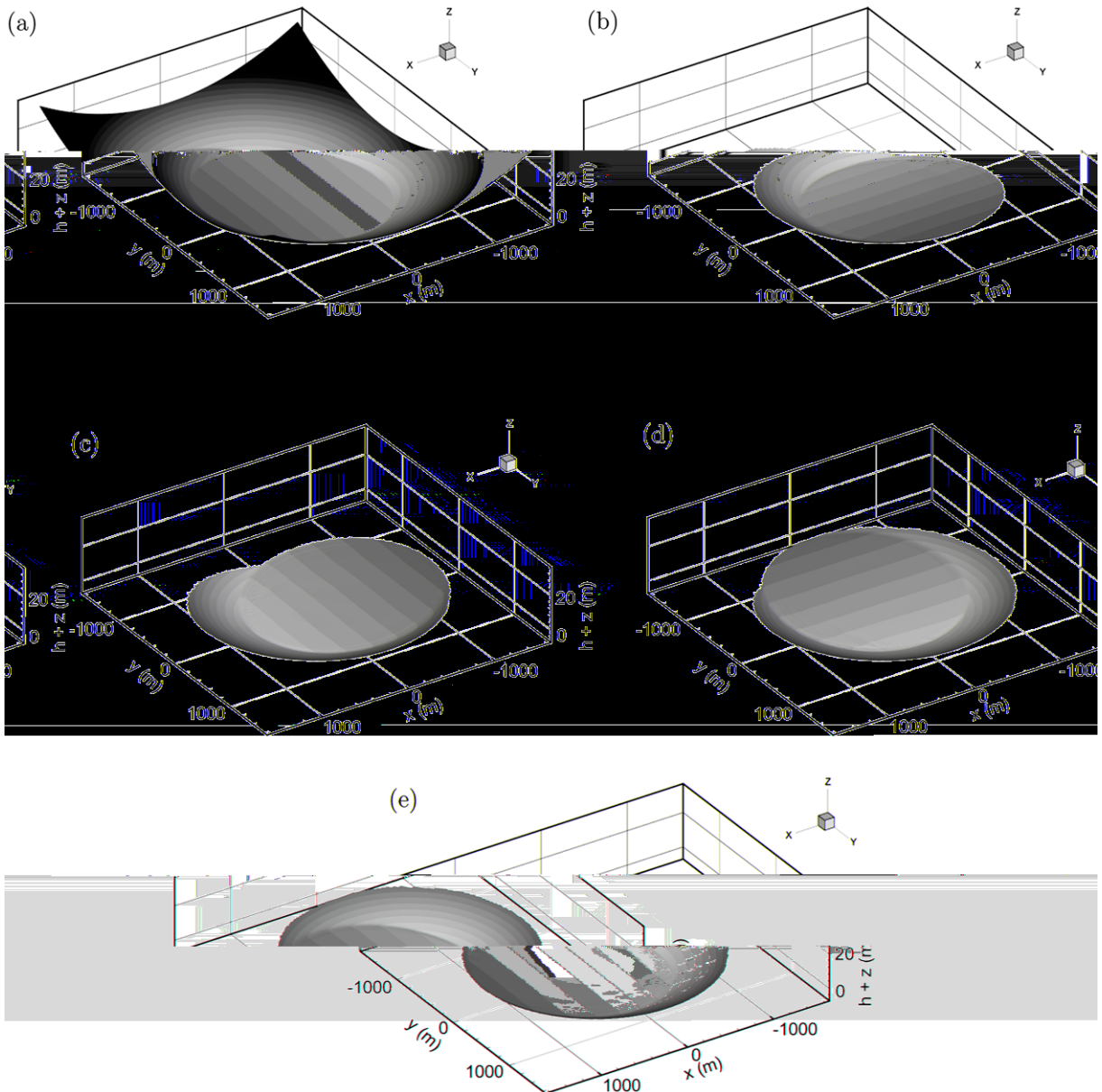


Fig. 26. Water level surface at (a) $t = 0$, (b) $t = T/4$, (c) $t = T/2$, (d) $t = 3T/4$ and (e) at $t = T$.

$$\mathbf{E}_k^* \mathbf{n}_k = \frac{1}{2} ((\mathbf{E}_i + \mathbf{E}_j) \mathbf{n} - \mathbf{P} |\tilde{\Lambda}| \mathbf{P}^{-1} \delta \mathbf{U})_k \tag{141}$$

and

$$(\mathbf{Tn})_k^\pm = (\mathbf{P} \mathbf{I}^\pm \mathbf{P}^{-1}) (\mathbf{Tn})_k \tag{142}$$

with $\mathbf{I}^\pm = \Lambda^{-1} \frac{1}{2} (\Lambda \pm |\Lambda|)$.

When cell averaging the solution in the 1D dimensional case the time step Δt is taken small enough so that there is no interaction of waves from neighboring Riemann problems, attending to a distance $\Delta x/2$. In the 2D framework, considering unstructured meshes, the equivalent distance to Δx , that will be referred to as χ_i in each cell i must consider the volume of the cell and the length of the shared k edges

$$\chi_i = \frac{A_i}{\max_{k=1,NE} l_k} \tag{143}$$

Considering that each k RP is used to deliver information between each pair of neighboring cells of different size, the associated distance $\min(A_i, A_j)/l_k$ is relevant, so in case that $\hat{h}(x', t) \geq 0$ in all k RP's the time step is limited by

$$\Delta t \leq CFL \Delta t^{\tilde{\lambda}}, \quad \Delta t^{\tilde{\lambda}} = \frac{\min(\chi_i, \chi_j)}{\max_{m=1,2,3} |\tilde{\lambda}^m|} \tag{144}$$

with $CFL = 1/2$ in the case of rectangular or structured triangular cells and, according to computational experience, CFL close to 1 for triangular unstructured grids as the construction of finite volume schemes from direct application of one-dimensional fluxes leads to reduced stability ranges [18].

It is possible to find values of $h < 0$ for some region of the approximate solution $\hat{\mathbf{U}}(x', t)$. In case of generating a negative value of h_j^{***} in a subcritical problem, as each RP is independent, to force stability, a limiting distance $\frac{1}{2} \chi_j$ is defined and the cell averaged scheme becomes:

$$h_j^{n+1} \frac{1}{2} \chi_j = h_j^{**} (\tilde{\lambda}_k^2 \Delta t) + h_j^{***} (\tilde{\lambda}_k^3 - \tilde{\lambda}_k^2) \Delta t + h_j^n \left(\frac{1}{2} \chi_j - \tilde{\lambda}_{i+1/2}^3 \right) \Delta t \geq 0. \tag{145}$$

Attending to the definition of $\tilde{\mathbf{e}}^2$ it is clear that $h_j^{**} = h_j^{***}$ and (145) is simplified leading to the following limit in the size of the time step

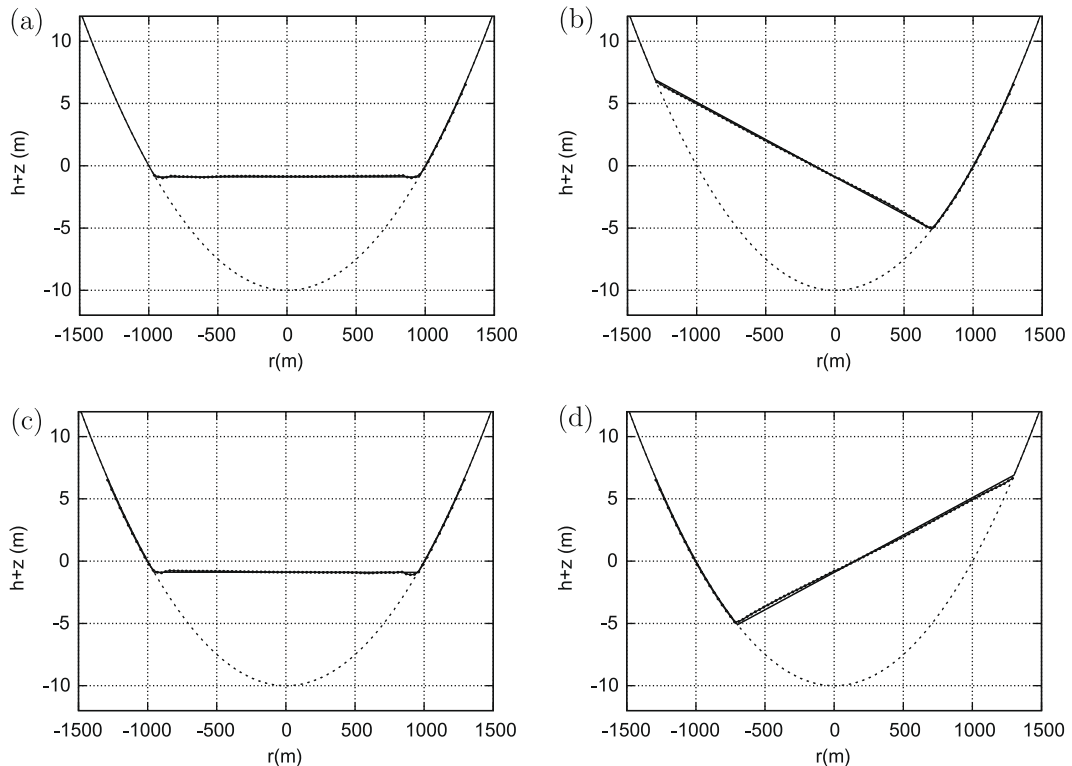


Fig. 27. Exact solution (–) and computed solution (– • –) for the water level surface at (a) $t = T/4$, (b) $t = T/2$, (c) $t = 3T/4$ and (d) at $t = T$.

$$\Delta t \leq \Delta t^{***}, \quad \Delta t^{***} = \frac{\chi_j}{2\tilde{\lambda}_k^3} \frac{h_j^n}{h_j^n - h_j^{***}}. \quad (146)$$

If $h_i^* < 0$ in the same subcritical case, the reduction in the time step is:

$$\Delta t \leq \Delta t^*, \quad \Delta t^* = \frac{\chi_i}{2\tilde{\lambda}_k^3} \frac{h_i^n}{h_i^n - h_i^*}. \quad (147)$$

It is possible to find regions with negative values of water depth near wet/dry interfaces with discontinuous bed level. According to (146) the time step becomes nil in those cases. To ensure positivity and conservation in the solution for all cases, the Godunov's method for the linear case is formulated as follows:

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n - \sum_{k=1}^{NE} (\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_{i,k}^- \frac{\Delta t l_k}{A_i}, \quad (148)$$

where the fluxes in a general intercell edge k are computed as follows:

- If $h_j^n = 0$ and $h_j^{***} < 0$ set:

$$(\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_{i,k}^- = (\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_k \quad (\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_{j,k}^- = 0. \quad (149)$$

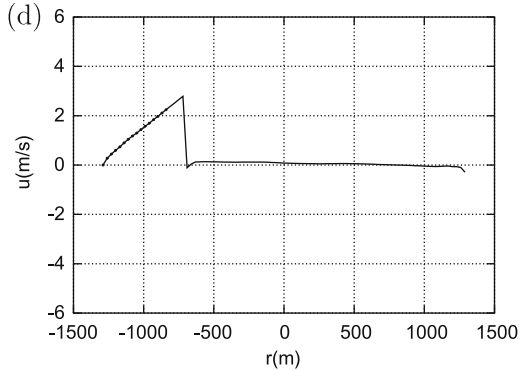
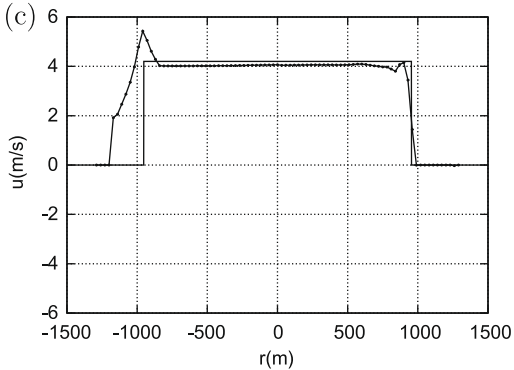
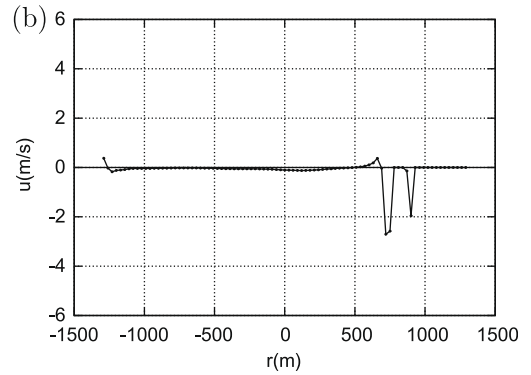
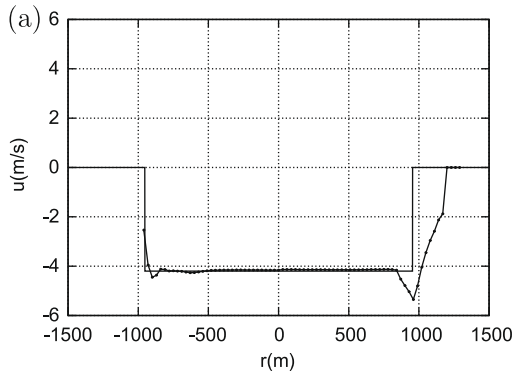
- If $h_i^n = 0$ and $h_i^* < 0$ set:

$$(\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_{j,k}^- = (\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_k \quad (\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_{i,k}^- = 0. \quad (150)$$

- Otherwise set:

$$(\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_{i,k}^- = \sum_{m=1}^{N_s} (\tilde{\lambda}^- \theta \alpha \tilde{\mathbf{e}}_k)^m, \quad (151)$$

$$(\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_{j,k}^- = (\delta \mathbf{E} \mathbf{n} - \mathbf{T} \mathbf{n})_{i,k}^+ = \sum_{m=1}^{N_s} (\tilde{\lambda}^+ \theta \alpha \tilde{\mathbf{e}}_k)^m.$$



Supercritical solutions as in (205) can lead to negative values of $h_i^* < 0$ or $h_j^{**} < 0$, although in these cases the cell averaging in the Godunov method avoids the appearance of negative values of h , as the source term does not participate in the updating of the water depth. In consequence the stability region becomes:

$$\Delta t \leq \begin{cases} \min(\Delta t^{***}, \Delta t^*, \Delta t^{\tilde{\lambda}}) & \text{if } (\tilde{\lambda}^1 \tilde{\lambda}^3)_k < 0. \\ \Delta t^{\tilde{\lambda}} & \text{otherwise,} \end{cases} \tag{152}$$

where Δt^{***} is defined as in (146) if $h_j^{**} < 0$ and $h_j^n \neq 0$ and Δt^* is defined as in (147) if $h_i^* < 0$ and $h_i^n \neq 0$.

To avoid unphysical results derived from the linearization, the version of the Harten–Hyman entropy fix [18] is applied. In the case of left transonic rarefaction $\lambda_i^1 < 0 < \lambda_j^1$, with $\lambda_i = \lambda(\mathbf{U}_i)$ and $\lambda_j = \lambda(\mathbf{U}_j)$, the initial jump associated to $\tilde{\lambda}_k^1$ is decomposed into two new jumps,

$$\begin{aligned} \bar{\lambda}_k^1 &= \lambda_i^1 \frac{(\lambda_j^1 - \tilde{\lambda}^1)}{(\lambda_j^1 - \lambda_i^1)}, & \hat{\lambda}_k^1 &= \lambda_j^1 \frac{(\tilde{\lambda}_k^1 - \lambda_i^1)}{(\lambda_j^1 - \lambda_i^1)}, \\ \bar{\beta}_k^1 &= \beta^1, & \hat{\beta}_k^1 &= 0, \end{aligned} \tag{153}$$

with $\bar{\lambda}_k^1 + \hat{\lambda}_k^1 = \tilde{\lambda}_k^1$ and $\bar{\beta}_k^1 + \hat{\beta}_k^1 = \beta_k^1$, preserving the original value of the state \mathbf{U}_i^* and, in consequence, the stability region in (152). For a right transonic rarefaction $\lambda_i^3 < 0 < \lambda_j^3$, the entropy fix procedure is entirely analogous and the single jump in $\tilde{\lambda}_k^3$ is split into two smaller jumps $\bar{\lambda}_k^3$ and $\hat{\lambda}_k^3$

$$\begin{aligned} \bar{\lambda}_k^3 &= \lambda_j^3 \frac{(\tilde{\lambda}_k^3 - \lambda_i^3)}{(\lambda_j^3 - \lambda_i^3)}, & \hat{\lambda}_k^3 &= \lambda_i^3 \frac{(\lambda_j^3 - \tilde{\lambda}_k^3)}{(\lambda_j^3 - \lambda_i^3)}, \\ \bar{\beta}_k^3 &= \beta^3, & \hat{\beta}_k^3 &= 0, \end{aligned} \tag{154}$$

with $\bar{\lambda}_k^3 + \hat{\lambda}_k^3 = \tilde{\lambda}_k^3$ and $\bar{\beta}_k^3 + \hat{\beta}_k^3 = \beta_k^3$, preserving the original value of the state \mathbf{U}_i^{**} and in consequence the stability region in (152).

4.3. Integration of the source term

Following the unified discretization in [5] the non-conservative term $(\mathbf{Tn})_k$ in (119) at a cell edge is written [13] as:

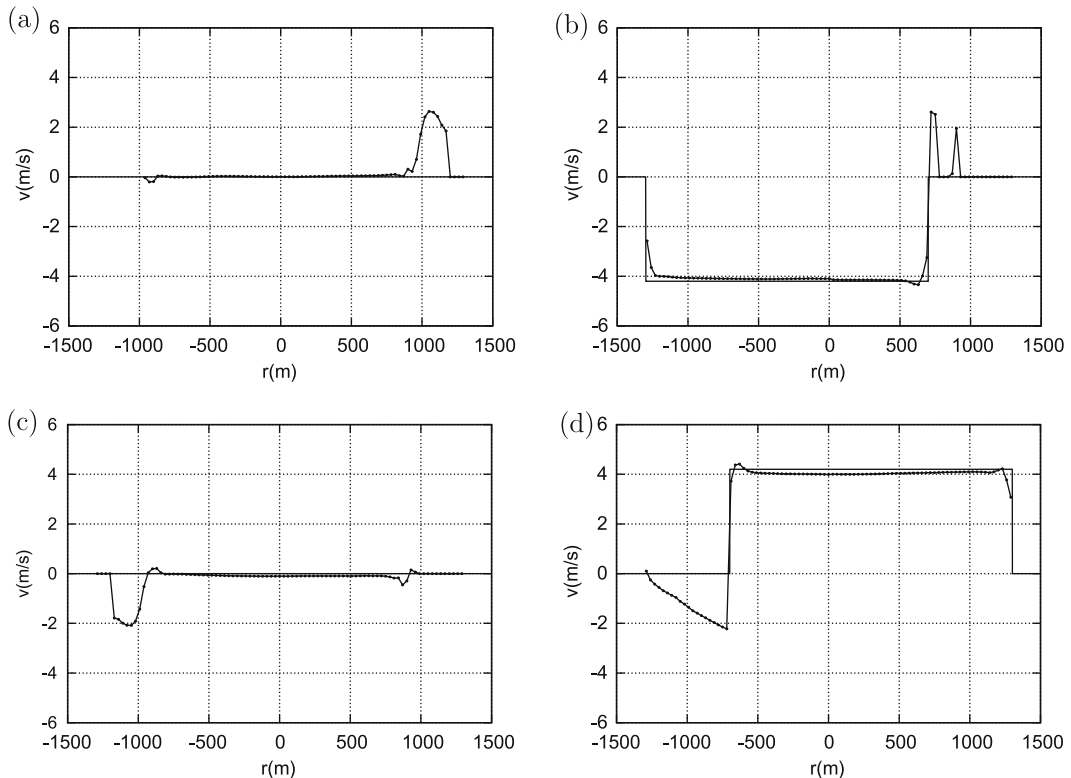


Fig. 29. Exact solution (—) and computed solution (— • —) for for velocity v at (a) $t = T/4$, (b) $t = T/2$, (c) $t = 3T/4$ and (d) at $t = T$.

$$(\mathbf{Tn})_k = \begin{pmatrix} 0 \\ \left(\frac{p_b}{\rho_w} - \frac{\tau_b}{\rho_w}\right) n_x \\ \left(\frac{p_b}{\rho_w} - \frac{\tau_b}{\rho_w}\right) n_y \end{pmatrix}_k, \tag{155}$$

where $\frac{p_b}{\rho_w}$ and $\frac{\tau_b}{\rho_w}$ attend to the pressure and friction exerted on the bed respectively.

In this work the following expression for the thrust term $\frac{p_b}{\rho_w}$ is proposed:

$$\left(\frac{p_b}{\rho_w}\right)_k = \begin{cases} \max\left(\left(\frac{p_b}{\rho_w}\right)_k^a, \left(\frac{p_b}{\rho_w}\right)_k^b\right) & \text{if } \delta d \delta z \geq 0 \text{ and } (\tilde{\mathbf{u}}\mathbf{n})\delta z > 0, \\ \left(\frac{p_b}{\rho_w}\right)_k^b & \text{otherwise,} \end{cases} \tag{156}$$

where $d = (h + z)$ and

$$\left(\frac{p_b}{\rho_w}\right)_k^a = -g(\tilde{h}\delta z)_k, \quad \left(\frac{p_b}{\rho_w}\right)_k^b = -g\left(h_r - \frac{|\delta z'|}{2}\right)\delta z', \tag{157}$$

with

$$r = \begin{cases} i & \text{if } \delta z \geq 0, \\ j & \text{if } \delta z < 0, \end{cases} \quad \delta z' = \begin{cases} h_i & \text{if } \delta z \geq 0 \text{ and } d_i < z_j, \\ h_j & \text{if } \delta z < 0 \text{ and } d_j < z_i, \\ \delta z & \text{otherwise.} \end{cases} \tag{158}$$

The discretization of the friction term based on [12] is applied

$$\left(\frac{\tau_b}{\rho_w}\right)_k = g(\tilde{h}S_f)_k d_n, \quad S_{f,k} = \left(\frac{n^2 \tilde{\mathbf{u}}\mathbf{n}|\tilde{\mathbf{u}}|}{\max(h_i, h_j)^{4/3}}\right)_k, \tag{159}$$

with d_n the normal distance between neighbor cell centers.

4.4. Reconstruction of the approximate solution $\hat{\mathbf{U}}(\mathbf{x}', t)$

The previous requirements on the properties of the numerical solution paying attention to the source terms can lead to extremely small values of Δt , as Δt^{***} or Δt^* can be various orders of magnitude smaller than $\Delta t^{\tilde{z}}$. This can be avoided by means of a reconstruction of the approximate solution $\hat{\mathbf{U}}(\mathbf{x}', t)$. The strategy proposed here is based on enforcing positive values of h_i^* and h_j^{**} when they become negative by reducing the numerical source term instead of reducing the time step size. Considering that $\tilde{e}_1^1 = 1$, positive values of h_i^* require that

$$h_i^* = h_i^n + \alpha_k^1 - \left(\frac{\beta}{\tilde{\lambda}}\right)_k^1 \geq 0 \tag{160}$$

leading to the following limit over β^1

$$\beta_k^1 \geq \beta_{min}^1, \quad \beta_{min}^1 = -(h_i^n + \alpha_k^1)|\tilde{\lambda}_k^1|. \tag{161}$$

As $\tilde{e}_1^2 = 0, h_j^{**} = h_j^{***}$ and considering that $\tilde{e}_1^3 = 1$, to obtain positive values of h_i^{***} it is necessary that

$$h_j^{***} = h_j^n - \alpha_k^3 + \left(\frac{\beta}{\tilde{\lambda}}\right)_k^3 \geq 0 \tag{162}$$

and a limit over β^3 appears

$$\beta_k^3 \geq \beta_{min}^3, \quad \beta_{min}^3 = -(h_j^n - \alpha_k^3)\tilde{\lambda}_k^3. \tag{163}$$

The reconstruction proposed in this work will be applied only to subcritical wet/wet RP, as in dry/wet RP the appearance of negative values of h_i^* or h_j^{***} in the approximate solution is helpful to provide a correct tracking of the flooding advance, and in supercritical cases positivity of the solution is guaranteed. Also cases where both $h_i^* < 0$ and $h_j^{***} < 0$ are omitted.

In the case that $h_i^* < 0$ and $h_j^{***} > 0$ and $\Delta t^* < \Delta t^{\tilde{z}}$, the new value of β^1 is redefined ensuring that h_j^{***} remains positive or null, as to ensure conservation β^3 must be replaced by the new value of $-\beta^1$, then

$$\beta^1 = \begin{cases} \beta_{min}^1 & \text{if } -\beta_{min}^1 \geq \beta_{min}^3 \quad \beta^3 = -\beta^1. \\ \beta^1 & \text{otherwise} \end{cases} \tag{164}$$

In the case that $h_i^* > 0$ and $h_j^{***} < 0$ and $\Delta t^{***} < \Delta t^{\tilde{z}}$, the new value of β^3 is redefined ensuring that h_i^* remains positive or null, as to ensure conservation β^1 must be replaced by the new value of $-\beta^3$, then

$$\beta^3 = \begin{cases} \beta_{min}^3 & \text{if } -\beta_{min}^3 \geq \beta^1 \\ \beta^3 & \text{otherwise} \end{cases} \quad \beta^1 = -\beta^3. \tag{165}$$

4.5. 2D application

The following test case is motivated by the necessity of checking the performance of reconstruction of the solution proposed in the previous section in two dimensional problems with transient boundaries dominated by the nonconservative terms and with analytical solution. Consider the frictionless parabolic topography defined by the depth function [17]

$$z(x,y) = -z_0 \left(1 - \frac{x^2 + y^2}{a^2} \right) \tag{166}$$

together with the periodic analytical solution of the two-dimensional shallow water equations stated in:

$$\begin{aligned} h(x,y,t) &= \max \left(0, \frac{\sigma z_0}{a^2} (2x \cos(\omega t) + 2y \sin(\omega t) - \sigma) - z(x,y) \right), \\ u(x,y,t) &= \sigma \omega \sin(\omega t), \quad v(x,y,t) = \sigma \omega \cos(\omega t), \end{aligned} \tag{167}$$

where u and v are the velocities in the x and y directions, and $\omega = \sqrt{2gz_0}/a$. Having shown the performance of the proposed scheme in cases with discontinuous source terms, this test case is useful to explore the limits of the approximated Riemann solver in two dimensional problems where the dry/wet boundaries are present.

The values $a = 1000$, $\sigma = 300$, $g = 9.80665$ and $z_0 = 10$ have been considered for this test, leading to a period $T = 448.57$ s. The computations have been performed using a triangular mesh built by drawing the diagonals on a quadrilateral grid with cell side $l = 15$ in a squared domain 3000×3000 m and using $CFL = 1.0$. Comparisons along a radius between the numerical and the analytical free surfaces at different times are shown in Figs. 26 and 27. The planar form of the free surface is maintained throughout the computation. In Figs. 28 and 29 comparisons for both the u and v velocities at time $t = T$ are shown.

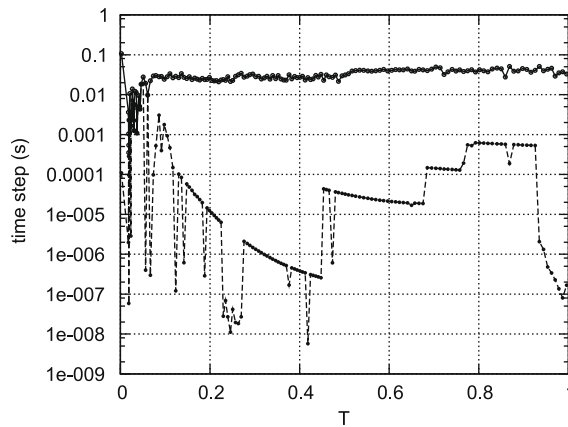


Fig. 30. Time step evolution in time for the original weak solution (—●—) and computed for the reconstructed weak solution (---○---).

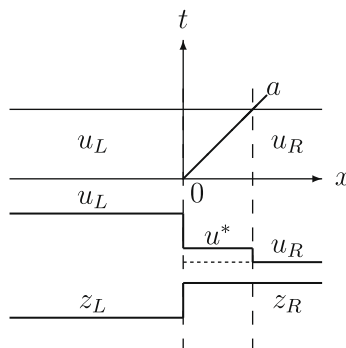


Fig. 31. Shock-wave solution of the Riemann problem.

The prediction of the velocities is good at all times in the cells covered by water according to the exact solution. The main discrepancies appear in the cells where the transient exchange of water is taking place with a certain lag time with respect to the exact solution. This example is representative of the time step size reduction imposed by the required positivity over the intermediate water depths in all the RP's. Figs. 30 and 31 shows the time history of the time step size automatically generated by the method following the conditions (152) (– • –). The positivity condition requires that the time step is reduced by three orders of magnitude in average along the total computational time. On the other hand, using the reconstruction based on limiting the discrete source term as in (164) and (164) enables a good quality solution at a constant time step size despite the transient flow.

5. Conclusions

In this paper new approximate Riemann solvers problems for equations and systems with source terms have been presented by adding one extra wave associated to the source term. The approximate solution is assumed discontinuous in $x = 0$ and therefore the flux cannot be correctly evaluated assuming the existence of a single value at that location. The method is specifically designed to satisfy the integral formulation. The method has not been designed specifically for source terms which provide some sort of balance with the flux derivatives so that the same techniques can easily be applied to all kind of source terms.

These wave solutions have been used to rebuild Roe's solver for the discretization of equations with source terms interpreted as non-conservative equations. The approach is based on the wave propagation algorithm in which waves arising in Riemann solutions are directly re-averaged onto adjacent cells in order to update the numerical solution. The method is applicable to both equations in the form of a conservation law as well as those where there is not a flux function (non-conservative equations). The careful procedure of wave averaging in the adjacent cells is the only way to control the stability region of the resulting method as well as the positivity of the variables in all kind of flow conditions. This is the starting point to explore different approximations to the integral source terms in the search for the best properties in each case.

The discretisation proposed is different from the work of many previous authors, who approximated their source terms in a manner which took into account the flux discretisation and, as a consequence, allowed the numerical model to maintain specific equilibria. Previous work dealing with the shallow water system has devoted particular attention to the special case of still water, and the schemes have been constructed so that they maintain this state. In this work, the emphasis is put on a more general idea that is based in the knowledge of non-trivial exact solutions. In fact, the improved accuracy of the new upwind discretisation of the source terms is shown in the approximation of other unsteady exact solutions, particularly in one dimension. Three approaches are discussed for the integral source term, all of them able to fit the trivial quiescent water steady state but different in other cases.

The effectiveness of these techniques has been illustrated, in the scalar case, for the Burgers' equation with source terms and also for systems of equations using the one- and two-dimensional shallow water equations, in which source terms are used to model variations in the bed topography and channel friction. This has been shown by comparison with a selection of test cases for which exact solutions are available.

Different methods of integrating source terms, which all respect the classical C-property for water at rest, have been shown to provide different behaviour in specific configurations when the water velocity is not equal to zero. Only one of them is recommended as it represents the best compromise in all kind of flows.

The work proposes two modifications over the original Roe's scheme: First, a reformulation of the stability condition that generalizes the classical CFL condition by including the influence of the source terms and the initial conditions in order to avoid the appearance of negative water depth values. Second, a redefinition of the numerical scheme in some particular cases in order to further prevent negative water depth values with independence of the time step. These modifications lead certainly to a considerable CPU time increase in complicated cases. It must be stressed, however, that the modifications are proposed precisely to make the basic method able to cope with complicated cases.

The careful design of the numerical scheme for the equations with source terms requires severe restriction of the allowable time step compatible with numerical stability and positivity in some cases. It is possible to avoid this by limiting the amount of numerical source term involved instead of limiting the time step and leading to more efficient computations. This part does not increase the CPU time but, on the contrary, is included to reduce it. It has been shown that, in combination with the source term approximation, the method produces accurate solutions for a wide variety of time-dependent test cases.

Acknowledgements

This work has been partially funded by the Spanish Ministry of Science and Technology under research project CGL2008-05153-C02-02.

Appendix A. Exact solution to Burgers' equation with source terms

Consider Burgers' equation including source terms as proposed in [11]:

$$\frac{\partial u}{\partial t} + \frac{1}{2} \frac{\partial u^2}{\partial x} = -u \frac{\partial z}{\partial x}, \quad (168)$$

with the initial data

$$u(x, 0) = u_0(x) = \begin{cases} u_L & \text{if } x < 0, \\ u_R & \text{if } x > 0, \end{cases} \quad z(x) = \begin{cases} z_L & \text{if } x < 0, \\ z_R & \text{if } x > 0. \end{cases} \quad (169)$$

A.1. Shock relations

Shock relations are moving discontinuities travelling at speed a that separate two general constant states, u_L^* and u_R^* [18,9]. In the homogeneous case they occur always when $u_L^* > u_R^*$ if a is positive.

Assuming that $u_L \geq 0$ and $u_R \geq 0$ there is a weak solution consisting of a steady discontinuity at $x=0$ plus a linear shock traveling with velocity a , connecting an intermediate state u^* with u_R :

$$u(x, t) = \begin{cases} u_L & \text{if } x < 0, \\ u^* & \text{if } 0 < x < at, \\ u_R & \text{if } x > at. \end{cases} \quad (170)$$

The presence of a source term in (168) leads to a standing discontinuity in $x=0$. The value of u^* can be derived from the steady state solution of Eq. (168):

$$\frac{du}{dx} = -\frac{dz}{dx} \Rightarrow du = -dz. \quad (171)$$

Integrating (171) between $x=0^+$ and $x=0^-$ the value of u^* is

$$u^* = u_L \zeta, \quad \zeta = \left(1 - \frac{\Delta z}{u_L}\right) \quad (172)$$

with $\Delta z = z_R - z_L$. The other discontinuity, moving with a shock speed a , must satisfy the Rankine–Hugoniot jump condition [18,9]:

$$\frac{1}{2}(u_R^2 - u^{*2}) = a(u_R - u^*) \quad (173)$$

leading to a celerity equal to:

$$a = \frac{1}{2}(u_R + u^*). \quad (174)$$

The parameter ζ measures the relative influence of the source term over the convective term and changes the solution of the RP. There are situations in which the weak solution is not unique. The entropy condition determines the physically relevant solutions. In solution (170) it is necessary that:

$$\lambda(u^*) > a > \lambda(u_R) \quad (175)$$

and as for (168) $\lambda(u) = u$, the solution is valid if:

$$\zeta > \frac{u_R}{u_L}. \quad (176)$$

In the particular case $\zeta = \frac{u_R}{u_L}$ the initial RP does not change in time.

A.2. Rarefaction waves

Rarefaction waves are weak solutions of the initial RP. Assuming that $u_R \geq 0$ and $u_L \geq 0$, the weak self-similar solution consists of a steady discontinuity at $x=0$ plus a right rarefaction wave expanding with a velocity x/t between two constant states u_L^* and u_R :

$$u(x, t) = \begin{cases} u_L & \text{if } x < 0, \\ u_L^* & \text{if } 0 < x < u_L^* t, \\ u^* = x/t & \text{if } u_L^* t < x < u_R t, \\ u_R & \text{if } x > u_R t. \end{cases} \quad (177)$$

The value of u_L^* can be derived considering the steady discontinuity in $x = 0$, as in (171)

$$u_L^* = u_L \zeta, \quad \zeta = \left(1 - \frac{\Delta z}{u_L}\right). \tag{178}$$

The solution is only acceptable if the following entropy condition holds (see Fig. 32):

$$\lambda(u_L^*) < \lambda(u^*) < \lambda(u_R) \tag{179}$$

and it is easy to check that the following limit is necessary to have a correct rarefaction solution

$$\zeta \leq \frac{u_R}{u_L}. \tag{180}$$

In the case $\zeta = \frac{u_R}{u_L}$ the solution in time is equal to the initial RP.

Considering that our RP is defined assuming $u_R \geq 0$ and $u_L \geq 0$, a positive solution characterized by values of $u(x, t) \geq 0$ is enforced if

$$0 \leq \zeta < \frac{u_R}{u_L}. \tag{181}$$

Appendix B. A three wave approximate Riemann solution

Depending on the flow conditions, three approximate solutions that satisfy (49), are proposed. The solutions for $\hat{\mathbf{U}}(x, t)$ are governed by the celerities in $\tilde{\lambda}_{i+1/2}$ and each one consists of four regions.

In the subcritical case, Fig. 33, the solution is given by:

$$\hat{\mathbf{U}}(x, t) = \begin{cases} \mathbf{U}_i^n & \text{if } x - \tilde{\lambda}^1 t < 0, \\ \mathbf{U}_i^* & \text{if } x - \tilde{\lambda}^1 t > 0 \text{ and } x < 0, \\ \mathbf{U}_{i+1}^{**} & \text{if } x > 0 \text{ and } x - \tilde{\lambda}^2 t < 0, \\ \mathbf{U}_{i+1}^n & \text{if } x - \tilde{\lambda}^2 t > 0 \end{cases} \tag{182}$$

with

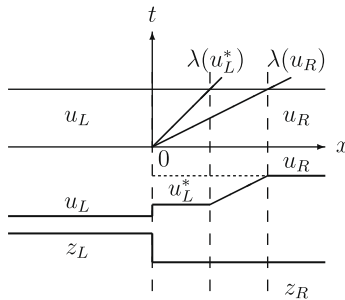


Fig. 32. Rarefaction wave solution of the Riemann problem.

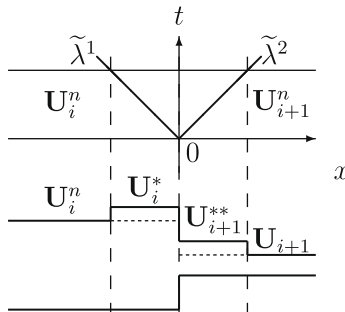


Fig. 33. Values of the solution \mathbf{U} in each wedge of the (x, t) plane for the subcritical case.

$$\begin{aligned} \mathbf{U}_i^*(\mathbf{U}_{i+1}, \mathbf{U}_i, \mathbf{S}_{i+1/2}) &= \mathbf{U}_i^n + (\theta\alpha\tilde{\mathbf{e}})_{i+1/2}^1, \\ \mathbf{U}_{i+1}^{**}(\mathbf{U}_{i+1}, \mathbf{U}_i, \mathbf{S}_{i+1/2}) &= \mathbf{U}_{i+1}^n - (\theta\alpha\tilde{\mathbf{e}})_{i+1/2}^2 \end{aligned} \tag{183}$$

characterized by

$$\begin{aligned} h_i^* &\neq h_{i+1}^{**}, \\ (hu)_i^* &= (hu)_{i+1}^{**} = \frac{\tilde{\lambda}^2\tilde{\lambda}^1(h_{i+1} - h_i) + \tilde{\lambda}^2(hu)_i - \tilde{\lambda}^1(hu)_{i+1}}{\tilde{\lambda}^2 - \tilde{\lambda}^1} + \beta^2 \end{aligned} \tag{184}$$

That is, the inner solutions in (182) conserve the same value of hu^* from left to right, and from right to left, but this symmetry disappears for h^* .

In the supercritical case, considering $\tilde{u} > 0$ the solution $\hat{\mathbf{U}}(x, t)$, Figs. 34 and 35, is given by:

$$\hat{\mathbf{U}}(x, t) = \begin{cases} \mathbf{U}_i^n & \text{if } x < 0, \\ \mathbf{U}_{i+1}^* & \text{if } x > 0 \text{ and } x - \tilde{\lambda}^1 t < 0, \\ \mathbf{U}_{i+1}^{**} & \text{if } x - \tilde{\lambda}^1 t > 0 \text{ and } x - \tilde{\lambda}^2 t < 0, \\ \mathbf{U}_{i+1}^n & \text{if } x - \tilde{\lambda}^2 t > 0, \end{cases} \tag{185}$$

with \mathbf{U}_{i+1}^{**} defined as in (183) and

$$\mathbf{U}_{i+1}^*(\mathbf{U}_{i+1}, \mathbf{U}_i, \mathbf{S}_{i+1/2}) = \mathbf{U}_{i+1}^{**} - (\theta\alpha\tilde{\mathbf{e}})_{i+1/2}^1 \tag{186}$$

characterized by

$$\begin{aligned} h_i^n &\neq h_{i+1}^*, \\ (hu)_i^n &= (hu)_{i+1}^* \end{aligned} \tag{187}$$

In the supercritical case, considering $\tilde{u} < 0$ the solution $\hat{\mathbf{U}}(x, t)$, is given by:

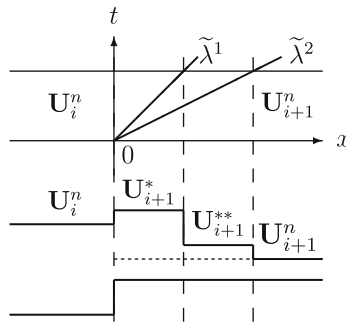


Fig. 34. Values of the solution \mathbf{U} in each wedge of the (x, t) plane for the supercritical case, $\tilde{u} > 0$.

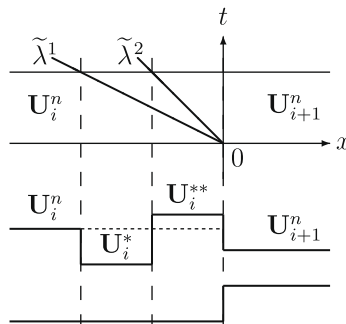


Fig. 35. Values of the solution \mathbf{U} in each wedge of the (x, t) plane for the supercritical case, $\tilde{u} < 0$.

$$\widehat{\mathbf{U}}(x, t) = \begin{cases} \mathbf{U}_i^n & \text{if } x - \tilde{\lambda}^1 t < 0 \\ \mathbf{U}_i^* & \text{if } x - \tilde{\lambda}^1 t > 0 \text{ and } x - \tilde{\lambda}^2 t < 0 \\ \mathbf{U}_i^{**} & \text{if } x - \tilde{\lambda}^2 t > 0 \text{ and } x < 0 \\ \mathbf{U}_{i+1}^n & \text{if } x > 0 \end{cases} \quad (188)$$

with \mathbf{U}_i^* defined as in (182) and

$$\mathbf{U}_i^{**}(\mathbf{U}_{i+1}, \mathbf{U}_i, \mathbf{S}_{i+1/2}) = \mathbf{U}_i^* + (\theta \alpha \tilde{\mathbf{e}})_{i+1/2}^2 \quad (189)$$

characterized by

$$\begin{aligned} h_i^{**} &\neq h_{i+1}^n \\ (hu)_i^{**} &= (hu)_{i+1}^n \end{aligned} \quad (190)$$

Appendix C. Dambreak over a bottom step

Letting aside considerations regarding the friction term, it is useful to study the effect of the bed slope in the solution of a dam break wave over a bottom step, where the bed discontinuity behaves as a stationary shock [2]. The initial data are given by

$$\begin{aligned} h(x, 0) &= h_o(x) = \begin{cases} h_L & \text{if } x < 0 \\ h_R & \text{if } x > 0 \end{cases} \\ u(x, 0) &= u_o(x) = 0 \\ z(x) &= \begin{cases} z_L & \text{if } x < 0 \\ z_R & \text{if } x > 0 \end{cases} \end{aligned} \quad (191)$$

The exact solution, as depicted in Fig. 36, consists of a left moving wave rarefaction, a steady jump in $x = 0$ and a right moving shock with a constant speed S_m . The rarefaction moves upstream at a velocity $\lambda^1(\mathbf{U}_L)$ connecting with a new state \mathbf{U}_L^* , that can be defined by means of the Riemann invariant across the wave

$$\frac{dh}{e_1^1} = \frac{d(hu)}{e_2^1} \quad (192)$$

and to be physically acceptable the following entropy relation which must be satisfied

$$\lambda^1(\mathbf{U}_L) < \lambda^1(\mathbf{U}_L^*) \quad (193)$$

The steady discontinuity at $x = 0$ connects two states, \mathbf{U}_L^* and \mathbf{U}_R^* . As temporal terms vanish

$$\frac{d\mathbf{F}}{dx} = \mathbf{S} \quad (194)$$

Integrating between $x = 0^+$ and $x = 0^-$ the value of \mathbf{U}_R^* is given by:

$$\begin{aligned} (hu)_R^* - (hu)_L^* &= 0 \\ \left(hu^2 + g \frac{h^2}{2} \right)_R^* - \left(hu^2 + g \frac{h^2}{2} \right)_L^* &= \int_{-0}^{+0} \frac{p_b}{\rho_w} dx \end{aligned} \quad (195)$$

The second equation in (195) involves the pressure on the vertical step in $x = 0$. Its evaluation actually requires to know the two-dimensional flow around the bottom step, which is another scientific topic. In a one-dimensional frame this closure-

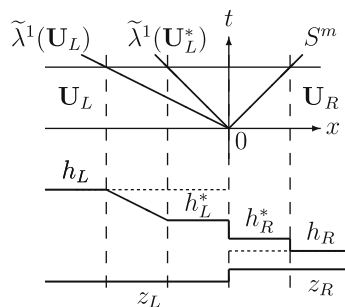


Fig. 36. Exact solution of the dam break problem.

problem may be faced by assuming that the pressure distribution is hydrostatic over the vertical step. The pressure term is [16]

$$\int_{-0}^{+0} \frac{p_b}{\rho_w} dx = -g \left(h_j - \frac{|z_R - z_L|}{2} \right) (z_R - z_L) \tag{196}$$

where

$$j = \begin{cases} L & \text{if } z_L \leq z_R \\ R & \text{if } z_L > z_R \end{cases} \tag{197}$$

According to [3] in order to exclude the multiplicity of solutions of (195), the entropy condition must be enforced; that is, total energy dissipates across the stationary shock.

The equations valid across the shock discontinuity propagating at a constant speed S_m , between constant states \mathbf{U}_R^* and \mathbf{U}_R can be derived from system (40). System of laws provide the following relation:

$$(\mathbf{F}(\mathbf{U}_R) - \mathbf{F}(\mathbf{U}_R^*)) = S_m (\mathbf{U}_R - \mathbf{U}_R^*) \tag{198}$$

and therefore

$$\begin{aligned} (hu)_R - (hu)_R^* &= S_m (h_R - h_R^*) \\ \left(hu^2 + g \frac{h^2}{2} \right)_R - \left(hu^2 + g \frac{h^2}{2} \right)_R^* &= S_m ((hu)_R - (hu)_R^*) \end{aligned} \tag{199}$$

A final remark concerns the entropy relation which must be satisfied in the case of a right moving shock in order to be physically acceptable:

$$\lambda^2(\mathbf{U}_R^*) > S_m > \lambda^2(\mathbf{U}_R) \tag{200}$$

Appendix D. A four wave approximate Riemann solution

Four approximate solutions that satisfy (122) are proposed. These solutions for $\hat{\mathbf{U}}(x', t)$ are governed by the celerities in $\tilde{\Lambda}_k$ and each one consists of five regions connected by four waves, one of them steady.

In the subcritical case with $\tilde{\mathbf{u}}\mathbf{n} > 0$, the proposed weak solution is:

$$\hat{\mathbf{U}}(x', t) = \begin{cases} \mathbf{U}_i^n & \text{if } x' - \tilde{\lambda}^1 t < 0 \\ \mathbf{U}_i^* & \text{if } x' - \tilde{\lambda}^1 t > 0 \text{ and } x' < 0 \\ \mathbf{U}_j^{**} & \text{if } x' > 0 \text{ and } x' - \tilde{\lambda}^2 t < 0 \\ \mathbf{U}_j^{***} & \text{if } x' - \tilde{\lambda}^2 t > 0 \text{ and } x' - \tilde{\lambda}^3 t < 0 \\ \mathbf{U}_j^n & \text{if } x' - \tilde{\lambda}^3 t > 0 \end{cases} \tag{201}$$

with

$$\begin{aligned} \mathbf{U}_i^* (\mathbf{U}_j, \mathbf{U}_i, (\mathbf{Tn})_k) &= \mathbf{U}_i^n + (\theta \alpha \tilde{\mathbf{e}})_k^1 \\ \mathbf{U}_j^{***} (\mathbf{U}_j, \mathbf{U}_i, (\mathbf{Tn})_k) &= \mathbf{U}_j^n - (\theta \alpha \tilde{\mathbf{e}})_k^3 \\ \mathbf{U}_j^{**} (\mathbf{U}_j, \mathbf{U}_i, (\mathbf{Tn})_k) &= \mathbf{U}_j^{***} - (\theta \alpha \tilde{\mathbf{e}})_k^2 \end{aligned} \tag{202}$$

as depicted in Figs. 37 and 38.

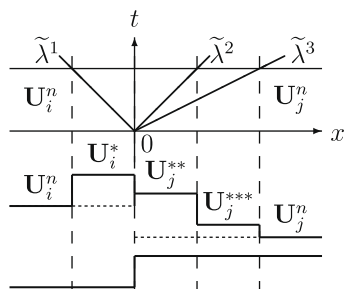


Fig. 37. Values of the solution \mathbf{U} in each wedge of the (x', t) plane for the subcritical case, $\tilde{\mathbf{u}}\mathbf{n} > 0$.

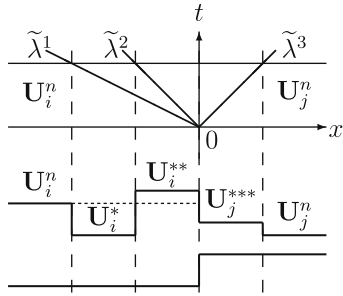


Fig. 38. Values of the solution \mathbf{U} in each wedge of the (x', t) plane for the subcritical case, $\tilde{\mathbf{u}}\mathbf{n} < 0$.

In the subcritical case, considering $\tilde{\mathbf{u}}\mathbf{n} < 0$, the proposed weak solution is:

$$\hat{\mathbf{U}}(x', t) = \begin{cases} \mathbf{U}_i^n & \text{if } x' - \tilde{\lambda}^1 t < 0 \\ \mathbf{U}_i^* & \text{if } x' - \tilde{\lambda}^1 t > 0 \text{ and } x' - \tilde{\lambda}^2 t < 0 \\ \mathbf{U}_i^{**} & \text{if } x' - \tilde{\lambda}^2 t < 0 \text{ and } x' < 0 \\ \mathbf{U}_j^{***} & \text{if } x' > 0 \text{ and } x' - \tilde{\lambda}^3 t < 0 \\ \mathbf{U}_j^n & \text{if } x' - \tilde{\lambda}^3 t > 0 \end{cases} \quad (203)$$

with

$$\begin{aligned} \mathbf{U}_i^*(\mathbf{U}_j, \mathbf{U}_i, (\mathbf{Tn})_k) &= \mathbf{U}_i^n + (\theta\alpha\tilde{\mathbf{e}})_k^1 \\ \mathbf{U}_i^{**}(\mathbf{U}_j, \mathbf{U}_i, (\mathbf{Tn})_k) &= \mathbf{U}_i^* + (\theta\alpha\tilde{\mathbf{e}})_k^2 \\ \mathbf{U}_j^{***}(\mathbf{U}_j, \mathbf{U}_i, (\mathbf{Tn})_k) &= \mathbf{U}_j^n - (\theta\alpha\tilde{\mathbf{e}})_k^3 \end{aligned} \quad (204)$$

In the supercritical case, considering $\tilde{\mathbf{u}}\mathbf{n} > 0$, is:

$$\hat{\mathbf{U}}(x', t) = \begin{cases} \mathbf{U}_i^n & \text{if } x' < 0 \\ \mathbf{U}_i^* & \text{if } x' > 0 \text{ and } x' - \tilde{\lambda}^1 t < 0 \\ \mathbf{U}_j^{**} & \text{if } x' - \tilde{\lambda}^1 t > 0 \text{ and } x' - \tilde{\lambda}^2 t < 0 \\ \mathbf{U}_j^{***} & \text{if } x' - \tilde{\lambda}^2 t > 0 \text{ and } x' - \tilde{\lambda}^3 t < 0 \\ \mathbf{U}_j^n & \text{if } x' - \tilde{\lambda}^3 t > 0 \end{cases} \quad (205)$$

and, in the supercritical case, assuming $\tilde{\mathbf{u}}\mathbf{n} < 0$, is:

$$\hat{\mathbf{U}}(x', t) = \begin{cases} \mathbf{U}_i^n & \text{if } x' - \tilde{\lambda}^1 t < 0 \\ \mathbf{U}_i^* & \text{if } x' - \tilde{\lambda}^1 t > 0 \text{ and } x' - \tilde{\lambda}^2 t < 0 \\ \mathbf{U}_i^{**} & \text{if } x' - \tilde{\lambda}^2 t > 0 \text{ and } x' - \tilde{\lambda}^3 t < 0 \\ \mathbf{U}_j^{***} & \text{if } x' - \tilde{\lambda}^3 t > 0 \text{ and } x' < 0 \\ \mathbf{U}_j^n & \text{if } x' > 0 \end{cases} \quad (206)$$

References

- [1] A.A. Akanbi, N.D. Katopodes, Model for flood propagation on initially dry land, *J. Hydraul. Eng.* 114 (1987) 689–706.
- [2] F. Alcrudo, F. Benkhalidoun, Exact solutions to the Riemann problem of the shallow water equations with a bottom step, *Comput. Fluids* 30 (2001) 643671.
- [3] R. Berneti, V.A. Titarev, E.F. Toro, Exact solution of the Riemann problem for the shallow water equations with discontinuous bottom geometry, *J. Comput. Phys.* 227 (2008) 3212–3243.
- [4] P. Brufau, M.E. Vázquez-Cendón, P. García-Navarro, A numerical model for the flooding and drying of irregular domains, *J. Numer. Methods Fluids* 39 (2002) 247–275.
- [5] J. Burguete, P. García-Navarro, Efficient construction of high-resolution TVD conservative schemes for equations with source terms: application to shallow water flows, *Int. J. Numer. Methods Fluids* 37 (2001) 209–248.
- [6] R. Courant, E. Isaacson, M. Rees, On the solution of nonlinear hyperbolic differential equations by finite differences, *Comm. Pure Appl. Maths.* 5 (1952) 243–255.
- [7] D.L. George, Augmented Riemann solvers for the shallow water equations over variable topography with steady states and inundation, *J. Comput. Phys.* 227 (2008) 3089–3113.
- [8] M.E. Hubbard, P. García-Navarro, Flux difference splitting and the balancing of source terms and flux gradients, *J. Comput. Phys.* 165 (2000) 89–125.
- [9] R.J. Leveque, *Finite Volume Methods for Hyperbolic Problems*, Cambridge University Press, New York, 2002. p. 311.

- [10] Q.H. Liang, F. Marche, Numerical resolution of well-balanced shallow water equations with complex source terms, *Adv. Water Resour.* 32 (2009) 873–884.
- [11] J. Murillo, J. Burguete, P. Brufau, P. García-Navarro, The influence of source terms on stability, accuracy and conservation in two-dimensional shallow flow simulation using triangular finite volumes, *Int. J. Numer. Methods Fluids* 54 (2007) 543–590.
- [12] J. Murillo, P. García-Navarro, J. Burguete, Time step restrictions for well balanced shallow water solutions in non-zero velocity steady states, *Int. J. Numer. Methods Fluids* 56 (2008) 661–686.
- [13] J. Murillo, P. García-Navarro, J. Burguete, Conservative numerical simulation of multicomponent transport in two-dimensional unsteady shallow water flow, *J. Comput. Phys.* 228 (2009) 5539–5573.
- [14] P.L. Roe, A basis for upwind differencing of the two-dimensional unsteady Euler equations, *Numerical Methods in Fluid Dynamics*, vol. II, Oxford University Press, Oxford, 1986.
- [15] B.D. Rogers, A.G.L. Borthwick, P.H. Taylor, Mathematical balancing of flux gradient and source terms prior to using Roes approximate Riemann solver, *J. Comput. Phys.* 168 (2003) 422–451.
- [16] G. Rosatti, J. Murillo, L. Fracrarollo, Generalized Roe schemes for 1D two-phase, free-surface flows over a mobile bed, *J. Comput. Phys.* 54 (2007) 543–590.
- [17] W.C. Thacker, Some exact solutions to the nonlinear shallow water equations, *J. Fluid Mech.* 107 (1981) 499–508.
- [18] E.F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer, Berlin, 1997. p. 526.
- [19] E.F. Toro, *Shock-Capturing Methods for Free-Surface Shallow Flows*, Wiley, New York, 2001. p. 109.
- [20] M.E. Vázquez-Cendón, Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry, *J. Comput. Phys.* 148 (1999) 497–498.
- [21] J.G. Zhou, D.M. Causon, C.G. Mingham, D.M. Ingram, The surface gradient method for the treatment of source terms in the shallow-water equations, *J. Comput. Phys.* 168 (2001) 1–25.