INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2003; **42**:23–55 (DOI: 10.1002/fld.436)

A flux-splitting solver for shallow water equations with source terms

Tomás Chacón Rebollo[*,1,†], Enrique D. Fernández Nieto[2,‡] and Macarena Gómez Mármol[3,§]

[1] *Departamento de Ecuaciones Diferenciales y Análisis Numérico, Universidad de Sevilla, C/Tarfia, s/n. 41080 Sevilla, Spain*
[2] *Departamento de Matemática Aplicada I, Universidad de Sevilla, E.T.S. Arquitectura, Avda, Reina Mercedes, s/n. 41080 Sevilla, Spain*
[3] *Departamento de Ecuaciones Diferenciales y Análisis Numérico, Universidad de Sevilla, C/Tarfia, s/n. 41080 Sevilla, Spain*

SUMMARY

This paper introduces a stable flux-splitting solver for one-dimensional (1D) shallow water equations. This solver is specifically designed to satisfy a strengthened consistency condition for stationary solutions that ensures the stability and accuracy of the scheme. It applies to channels with variable depth and width, including terms modelling friction at bottom and vertical walls. Some numerical tests by comparison to both analytical solutions and experimental measurements show the good performances of the scheme. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS: finite-volume method; upwinding; shallow water; flux-splitting; source terms

1. INTRODUCTION

This paper deals with flux-splitting numerical solvers for one-dimensional (1D) shallow water equations with source terms.

We shall consider shallow water flows in straight channels with variable depth and width, in the presence of friction effects. We assume that this flow is governed by the 1D shallow

* Correspondence to: T. Chacón Rebollo, Departamento de Ecuaciones Diferenciales y Análisis Numérico, Universidad de Sevilla, C/Tarfia, s/n. 41080 Sevilla, Spain.
† E-mail: chacon@numer.us.es
‡ E-mail: edofer@us.es
§ E-mail: macarena@numer.us.es

Contract/grant sponsor: Spanish Government Research Project; contract/grant number: REN2000-1162-C02-01
Contract/grant sponsor: Spanish Government Research Project; contract/grant number: REN2000-1168-C02-01

Received 10 November 2001
Revised 11 September 2002

water equations

$$\frac{\partial}{\partial t}W + \frac{\partial}{\partial x}F(W) = G(x, W) \quad \text{in }]0, L[\times]0, T[\tag{1}$$

$$W = W(x,t) = \begin{pmatrix} h(x,t) \\ q(x,t) \end{pmatrix} \tag{2}$$

where $h(x,t)$ and $q(x,t)$, respectively, denote the height of the water column and the discharge across the channel section at position $x \in [0, L]$ and time $t \in [0, T]$. F denotes the physical flux, given by

$$F(W) = \begin{pmatrix} q \\ \dfrac{q^2}{h} + \dfrac{1}{2}gh^2 \end{pmatrix} \tag{3}$$

and G is the source term, which may take into account variable depth and width of the channel, and also friction effects. The precise definition of G shall be given later on.

Equation (1) should be completed with initial and suitable boundary conditions in order to have a well-posed problem.

We shall consider finite-volume numerical solvers. Given a mesh $\{x_i\}_{i=0}^{N+1}$, and a time step Δt, we shall approximate the solution $W(x,t)$ of (1) by a piecewise constant function that takes the constant value W_i^n on the cell $]x_{i-1/2}, x_{i+1/2}[\times]t_n, t_{n+1}[$, where $x_{i+1/2} = (x_i + x_{i+1})/2$ and $t_n = n\Delta t$. The values $\{W_i^n\}_{n,i}$ are obtained as the solution of general explicit solvers of the form

$$\frac{W_i^{n+1} - W_i^n}{\Delta t} + \frac{\phi(W_i^n, W_{i+1}^n) - \phi(W_{i-1}^n, W_i^n)}{\Delta x_i} = \mathcal{G}(x_{i-1}, x_i, x_{i+1}, W_{i-1}^n, W_i^n, W_{i+1}^n) \tag{4}$$

complemented with suitable initial and boundary conditions, and where

$$\Delta x_i = x_{i+1/2} - x_{i-1/2}$$

Here, the function $\phi(W_i, W_{i+1})$ is usually known as the *numerical flux*, while the function $\mathcal{G}(x_{i-1}, x_i, x_{i+1}, W_{i-1}, W_i, W_{i+1})$ is called the *numerical source term*.

In general, the presence of source terms in non-linear hyperbolic conservation laws sets specific stability and accuracy restrictions to numerical solvers. A stable and accurate discretization of such equations seems to strongly rely, in addition to the well-known upwinding of standard hyperbolic conservation laws, upon high enough spatial resolution (Cf. References [1–4, 15]). In the case of shallow water equations, this has been formalized by Bermúdez and Vázquez Cendón in Reference [5] as an enhanced consistency condition. We shall precisely define here this condition, as it is at the base of our analysis.

Definition 1 (*Bermúdez–Vázquez condition*)
Scheme (4) is said to verify the exact (respectively, approximated) Bermúdez–Vázquez consistency condition with respect to a stationary solution $\bar{W}$ of problem (1) if it exactly

(respectively, up to an error $O(\Delta x_i)^2$) solves this stationary solution at grid nodes, i.e.

$$\frac{1}{\Delta x}[\phi(\bar{W}_i, \bar{W}_{i+1}) - \phi(\bar{W}_{i-1}, \bar{W}_i)] = \mathcal{G}(x_{i-1}, x_i, x_{i+1}, \bar{W}_{i-1}, \bar{W}_i, \bar{W}_{i+1}) \quad (\text{resp.}, +O(\Delta x_i)^2) \tag{5}$$

$$\forall x_i \in]0, L[, \text{ where } \bar{W}_i = \bar{W}(x_i)$$

Although there is still no theoretical analysis of the role played by this condition, in practice it ensures stable and accurate computations for 1D and 2D flow models.

In Reference [6] the methods of Roe and Van Leer are extended to shallow water equations with source terms, in order to verify the Bermúdez and Vázquez condition. However, no extensions of flux-splitting methods, such as Steger–Warming's or Vijayasundaram's are reported. We recall that flux-splitting methods can be derived when the flux F verifies a property of homogeneity, concretely when $F(W) = A(W)W$. This is the case of, for example, the 1D or 2D shallow water equations as well as Euler equations.

Flux-splitting methods have properties different from flux-difference methods, as can be the behaviour in sonic points and shock waves. For example, the flux-splitting method of Vijayasundaram does not present problems of lack of viscosity in sonic points. This may make it preferable the use of flux-splitting methods in some situations.

We introduce in this paper a flux-splitting scheme that verifies the enhanced consistency condition. It is a mixing of Steger–Warming and Vijayasundaram schemes, the first being applied to the momentum conservation equation and the second one to the mass conservation equation. Methods such AUSM and the flux-splitting method of VanLeer (see Reference [7]) present a similar treatment, but for Euler equations. The objective is to distinguish the treatment of the equations that contains the terms of pressure.

Our main methodological innovation is based upon the following idea: the upwinding of the source term is usually performed via a projection onto the eigenvectors of the flux matrix. In our case, these eigenvectors are re-scaled by constants, which are determined by imposing the enhanced consistency condition. This gives more flexibility to our scheme, while preserving the correct upwinding directions on each characteristic field.

We obtain a scheme which exactly calculates a stationary solution for variable depth and width of the channel, and up to order two if friction terms are included. It performs quite closely to the extension of Roe's scheme, in fact both schemes yield very close solutions with practically the same computational effort.

Furthermore, the technique that is introduced for the treatment of source terms can be applied to Euler equations and to the generalization of other flux-splitting methods, for example, directly Steger–Warming's or Vijayasundaram's.

The paper is organized as follows. Section 2 introduces the main solver we are dealing with. Section 3 is devoted to the application of this solver to channels with variable depth and constant width, in the absence of friction effects. We analyse the enhanced consistency condition, and present some numerical tests for analytical solutions. In Section 4, we also prove that our scheme verifies the enhanced consistency condition for source terms including variable depth and width, and also friction terms. We test this general case by comparison with both analytic solutions and experimental measurements. Finally, in Section 5, we apply

our technique for extending flux-splitting solvers to a different solver, in order to test the flexibility of our technique.

2. NUMERICAL SOLVER

In this section, we introduce the numerical flux and the numerical source term for the solver we consider, for general source terms.

2.1. Numerical flux

Flux-splitting solvers are based upon the decomposition $F(W)=A(W)W$, where

$$A(W)=\begin{pmatrix} 0 & 1 \\ -\dfrac{q^2}{h^2}+\dfrac{1}{2}gh & 2\dfrac{q}{h} \end{pmatrix} \tag{6}$$

Matrix A is diagonalizable if $h>0$, concretely $A=X\Lambda X^{-1}$, where

$$\Lambda=\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad \lambda_1=\frac{q}{h}+\sqrt{\frac{1}{2}gh}, \quad \lambda_2=\frac{q}{h}-\sqrt{\frac{1}{2}gh} \tag{7}$$

and

$$X=\begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \tag{8}$$

This naturally yields the flux decomposition

$$F(W)=F^+(W)+F^-(W)$$

where

$$F^{\pm}(W)=(A)^{\pm}(W)W \quad \text{with } (A)^{\pm}(W)=X\Lambda^{\pm}X^{-1} \tag{9}$$

and

$$\Lambda^+=\begin{pmatrix} \max(\lambda_1,0) & 0 \\ 0 & \max(\lambda_2,0) \end{pmatrix}, \quad \Lambda^-=\begin{pmatrix} \min(\lambda_1,0) & 0 \\ 0 & \min(\lambda_2,0) \end{pmatrix} \tag{10}$$

In accordance with this decomposition, the flux-splitting schemes are built from numerical flux functions of the form

$$\phi(W_i,W_{i+1})=\phi^+(W_i,W_{i+1})+\phi^-(W_i,W_{i+1}) \tag{11}$$

with

$$\phi^+(W_i,W_{i+1})=B_1(W_i,W_{i+1})W_i, \quad \phi^-(W_i,W_{i+1})=B_2(W_i,W_{i+1})W_{i+1} \tag{12}$$

where B_1 and B_2 are 2×2 matrices that should be specified for each actual scheme. ϕ^+ and ϕ^-, respectively, represent upwinding to the left and to the right.

A consistent scheme for the homogeneous equation is obtained if B_1 and B_2 verify

$$F^+(W) = B_1(W, W)W, \quad F^-(W) = B_2(W, W)W \tag{13}$$

Our scheme is based upon Vijayasundaram (Cf. References [8, 9]) and Steger–Warming (Cf. References [7, 10]) schemes, which, respectively, correspond to

$$B_1(W_i, W_{i+1}) = A^+\left(\frac{W_i + W_{i+1}}{2}\right), \quad B_2(W_i, W_{i+1}) = A^-\left(\frac{W_i + W_{i+1}}{2}\right)$$

and

$$B_1(W_i, W_{i+1}) = A^+(W_i), \quad B_2(W_i, W_{i+1}) = A^-(W_{i+1})$$

We shall apply Vijayasundaram's scheme to the first equation in system (1) and Steger–Warming's scheme to the second one. Let us denote by $(a_1)^+$ (resp., $(a_2)^+$) the first (resp., second) row in matrix $(A)^+$, and similarly $(a_1)^-$ and $(a_2)^-$. We define our flux-splitting matrices as

$$B_1(W_i, W_{i+1}) = \begin{pmatrix} (a_1)^+((W_i + W_{i+1})/2) \\ \hline (a_2)^+(W_i) \end{pmatrix}, \quad B_2(W_i, W_{i+1}) = \begin{pmatrix} (a_1)^-((W_i + W_{i+1})/2) \\ \hline (a_2)^-(W_{i+1}) \end{pmatrix} \tag{14}$$

This corresponds to the following numerical flux:

$$\phi(W_i, W_{i+1}) = \begin{pmatrix} (a_1)^+((W_i + W_{i+1})/2) \\ \hline (a_2)^+(W_i) \end{pmatrix} W_i + \begin{pmatrix} (a_1)^-((W_i + W_{i+1})/2) \\ \hline (a_2)^-(W_{i+1}) \end{pmatrix} W_{i+1} \tag{15}$$

Notice that B_1 and B_2 verify (13), and consequently the scheme is consistent for the homogeneous equation.

2.2. Numerical source term

To define the numerical source term for our scheme, we shall give a generalization of the procedure defined in Bermúdez and Vázquez Cendón [6].

Let us consider a 'upwind' diagonalizable matrix $\tilde{A}$, which filters out negative eigenvalues of the flux matrix A. Following References [10, 11], we decompose the source term as

$$G(x_i, W_i) = \tfrac{1}{2}[I + |\tilde{A}|\tilde{A}^{-1}]G(x_i, W_i) + \tfrac{1}{2}[I - |\tilde{A}|\tilde{A}^{-1}]G(x, W) \tag{16}$$

where $|\tilde{A}|$ is a matrix whose eigenvalues are the absolute values of the eigenvalues of $\tilde{A}$ with the same eigenvectors. Thus, the first and second summands, respectively, correspond to positive and negative eigenvalues. This suggests to split the numerical source term as

$$\mathcal{G}(x_{i-1}, x_i, x_{i+1}, W_{i-1}, W_i, W_{i+1}) = \mathcal{G}_{\mathrm{L}}(x_{i-1}, x_i, W_{i-1}, W_i) + \mathcal{G}_{\mathrm{R}}(x_i, x_{i+1}, W_i, W_{i+1}) \tag{17}$$

with

$$\mathcal{G}_{\mathrm{L}}(x_{i-1},x_i,W_{i-1},W_i) = \frac{(x_i-x_{i-1})/2}{\Delta x_i}[I+|B(W_{i-1},W_i)|B^{-1}(W_{i-1},W_i)]\tilde{G}(x_{i-1},x_i,W_{i-1},W_i) \tag{18}$$

$$\mathcal{G}_{\mathrm{R}}(x_i,x_{i+1},W_i,W_{i+1}) = \frac{(x_{i+1}-x_i)/2}{\Delta x_i}[I-|B(W_i,W_{i+1})|B^{-1}(W_i,W_{i+1})]\tilde{G}(x_i,x_{i+1},W_i,W_{i+1}) \tag{19}$$

where $B(W_i,W_{i+1})$ is some approximation of $\tilde{A}((W_i+W_{i+1})/2)$ and $\tilde{G}(x_i,x_{i+1},W_i,W_{i+1})$ is some approximation of $G(x_{i+1/2},W_{i+1/2})$. Both $B(W_i,W_{i+1})$ and $\tilde{G}(x_i,x_{i+1},W_i,W_{i+1})$ should be provided to have a completely defined scheme.

In Reference [5], matrix $\tilde{A}$ is actually chosen as A. This yields extensions of Roe's and Van Leer schemes satisfying the enhanced consistency condition. However, this choice does not provide an extension of standard flux-splitting schemes satisfying such condition. Instead, we observe that the upwinding must be governed by the signs of the eigenvalues of A, while its eigenvectors play a less important role in the upwinding of the source term. We define

$$\tilde{A}=\tilde{X}\Lambda\tilde{X}^{-1} \quad \text{where } \Lambda=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix} \tag{20}$$

and $\tilde{X}=PX$, P being an 'eigenvectors scaling' matrix, defined as

$$P=\begin{pmatrix}c_1 & 0\\ 0 & c_2\end{pmatrix} \quad \text{with } c_1\neq 0,\ c_2\neq 0 \tag{21}$$

We then define $B(W_i,W_{i+1})$ as

$$B(W_i,W_{i+1})=\tilde{A}\left(\frac{W_i+W_{i+1}}{2}\right) \tag{22}$$

As we shall see, this introduces some flexibility in our flux-splitting scheme, which will verify the enhanced consistency condition for some choices of the scaling matrix P.

Notice that if $B=\tilde{A}$, the upwinding matrices B_1 and B_2 of the flux term defined in (14) and matrix B are related by

$$B=P(B_1+B_2)P^{-1}$$

if all are evaluated at the same point.

3. VARIABLE DEPTH

In this section, we analyse the enhanced consistency condition, and perform some numerical tests for our scheme, in the case of variable depth and constant width of the channel, without

considering friction terms. In this case, the source term in (1) reads

$$G(x,W)=\begin{pmatrix}0\\ ghH'(x)\end{pmatrix} \tag{23}$$

where $H(x)$ is the channel bottom profile function, so that the bottom equation is $z=-H(x)$. We assume that $H(x)$ is a positive and continuously differentiable function on $[0,L]$.

In this case, a stationary solution of Equations (1) is given by (Cf. Reference [5])

$$\bar{W}(x)=\begin{pmatrix}H(x)\\ 0\end{pmatrix} \tag{24}$$

We state the following result.

Theorem 1

When $c_1=1$, $c_2=2$, the flux-splitting scheme (4), (15), (17) through (22) satisfies the exact Bermúdez–Vázquez condition with respect to the stationary solution (24) if

$$\tilde{G}(x,y,(h_{\mathrm{L}},q_{\mathrm{L}}),(h_{\mathrm{R}},q_{\mathrm{R}}))=\begin{pmatrix}0\\ g\dfrac{h_{\mathrm{L}}+h_{\mathrm{R}}}{2}\dfrac{H(y)-H(x)}{y-x}\end{pmatrix} \tag{25}$$

This scheme also satisfies the approximated Bermúdez–Vázquez condition if

$$\tilde{G}(x,y,(h_{\mathrm{L}},q_{\mathrm{L}}),(h_{\mathrm{R}},q_{\mathrm{R}}))=\begin{pmatrix}0\\ g\dfrac{h_{\mathrm{L}}+h_{\mathrm{R}}}{2}H'\left(\dfrac{x+y}{2}\right)\end{pmatrix} \tag{26}$$

Proof

We start by proving that if $\tilde{G}$ is given by (25), then our scheme satisfies the enhanced consistency condition. We assume $c_1=1$ as in fact the only relevant constant is $c=c_2/c_1$.

For simplicity of notation, we shall denote $\bar{W}(x)$ by $\bar{W}$ and $\bar{W}(x_i)$ by $\bar{W}_i$. Using (7), (8) and (10) we deduce

$$(A)^{+}(\bar{W})=\frac{1}{2}\begin{pmatrix}\sqrt{\frac{1}{2}gh} & 1\\ \frac{1}{2}gh & \sqrt{\frac{1}{2}gh}\end{pmatrix}$$

$$(A)^{-}(\bar{W})=(A-A^{+})(\bar{W})=\frac{1}{2}\begin{pmatrix}-\sqrt{\frac{1}{2}gh} & 1\\ \frac{1}{2}gh & -\sqrt{\frac{1}{2}gh}\end{pmatrix}$$

Thus,

$$(a_1)^{+}(\bar{W})=\frac{1}{2}\left(\sqrt{\frac{1}{2}gh},1\right),\quad (a_2)^{+}(\bar{W})=\frac{1}{2}\left(\frac{1}{2}gh,\sqrt{\frac{1}{2}gh}\right)$$

and

$$(a_1)^-(\bar{W}) = \frac{1}{2}\left(-\sqrt{\frac{1}{2}gh}, 1\right), \quad (a_2)^-(\bar{W}) = \frac{1}{2}\left(\frac{1}{2}gh, -\sqrt{\frac{1}{2}gh}\right)$$

The numerical flux is then given by

$$\phi(\bar{W}_i, \bar{W}_{i+1}) = \left(\frac{(a_1)^+\left(\frac{h_i + h_{i+1}}{2}, 0\right)}{(a_2)^+(h_i, 0)}\right)\begin{pmatrix} h_i \\ 0 \end{pmatrix}$$

$$+ \left(\frac{(a_1)^-\left(\frac{h_i + h_{i+1}}{2}, 0\right)}{(a_2)^-(h_{i+1}, 0)}\right)\begin{pmatrix} h_{i+1} \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -\sqrt{\frac{1}{2}\, g\frac{h_i + h_{i+1}}{2}}(h_{i+1} - h_i) \\ \frac{1}{2}\, g(h_i^2 + h_{i+1}^2) \end{pmatrix} \tag{27}$$

Consequently,

$$\frac{1}{\Delta x_i}(\phi(\bar{W}_i, \bar{W}_{i+1}) - \phi(\bar{W}_{i-1}, \bar{W}_i))$$

$$= \frac{1}{2}\frac{1}{\Delta x_i}\begin{pmatrix} \frac{1}{2}\sqrt{g(h_{i-1} + h_i)}(h_i - h_{i-1}) - \frac{1}{2}\sqrt{g(h_i + h_{i+1})}(h_{i+1} - h_i) \\ \frac{1}{2}g(h_{i+1}^2 - h_{i-1}^2) \end{pmatrix} \tag{28}$$

We next calculate the numerical source. Using

$$\tilde{X}(\bar{W}) = \begin{pmatrix} 1 & 1 \\ c\lambda_1 & c\lambda_2 \end{pmatrix} \quad \text{with } \lambda_2 = -\lambda_1$$

we deduce

$$\tilde{A}(\bar{W}) = \begin{pmatrix} 0 & 1/c \\ c\lambda_1{}^2 & 0 \end{pmatrix}, \quad |\tilde{A}(\bar{W})| = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}$$

and

$$\tilde{A}^{-1}(\bar{W}) = \begin{pmatrix} 0 & 1/(c\lambda_1^2) \\ c & 0 \end{pmatrix}$$

where $\lambda_i = \lambda_i(\bar{W})$. From these expressions, we obtain

$$I + |\tilde{A}(\bar{W})|\tilde{A}^{-1}(\bar{W}) = \begin{pmatrix} 1 & \dfrac{1}{c\lambda_1} \\ c\lambda_1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \dfrac{1}{c\sqrt{\frac{1}{2}gh}} \\ c\sqrt{\frac{1}{2}gh} & 1 \end{pmatrix}$$

$$I - |\tilde{A}(\bar{W})|\tilde{A}^{-1}(\bar{W}) = \begin{pmatrix} 1 & -\dfrac{1}{c\lambda_1} \\ -c\lambda_1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \dfrac{-1}{c\sqrt{\frac{1}{2}gh}} \\ -c\sqrt{\frac{1}{2}gh} & 1 \end{pmatrix}$$

Now, using $B(\bar{W}_1, \bar{W}_2) = \tilde{A}((\bar{W}_1 + \bar{W}_2)/2)$, from Equations (18) and (19) we obtain the expressions for $\mathscr{G}_{\mathrm{L}}$ and $\mathscr{G}_{\mathrm{R}}$,

$$\begin{aligned} &\mathscr{G}_{\mathrm{L}}(x_{i-1}, x_i, \bar{W}_{i-1}, \bar{W}_i) \\ &= \frac{(x_i - x_{i-1})/2}{\Delta x_i} \begin{pmatrix} 1 & \dfrac{1}{c\sqrt{\frac{1}{2}\, g\, \frac{h_{i-1}+h_i}{2}}} \\ c\sqrt{\frac{1}{2}\, g\, \frac{h_{i-1}+h_i}{2}} & 1 \end{pmatrix} \begin{pmatrix} 0 \\ g\, \dfrac{h_{i-1}+h_i}{2}\, \dfrac{H(x_i) - H(x_{i-1})}{x_i - x_{i-1}} \end{pmatrix} \\ &= \frac{1}{2}\frac{1}{\Delta x_i} \begin{pmatrix} \dfrac{2}{c}\sqrt{\frac{1}{2}\, g\, \frac{h_{i-1}+h_i}{2}}\,(H(x_i) - H(x_{i-1})) \\ g\, \dfrac{h_{i-1}+h_i}{2}\,(H(x_i) - H(x_{i-1})) \end{pmatrix} \end{aligned} \tag{29}$$

$$\begin{aligned} &\mathscr{G}_{\mathrm{R}}(x_i, x_{i+1}, \bar{W}_i, \bar{W}_{i+1}) \\ &= \frac{(x_{i+1} - x_i)/2}{\Delta x_i} \begin{pmatrix} 1 & \dfrac{-1}{c\sqrt{\frac{1}{2}\, g\, \frac{h_i+h_{i+1}}{2}}} \\ -c\sqrt{\frac{1}{2}\, g\, \frac{h_i+h_{i+1}}{2}} & 1 \end{pmatrix} \begin{pmatrix} 0 \\ g\, \dfrac{h_i+h_{i+1}}{2}\, \dfrac{H(x_{i+1}) - H(x_i)}{x_{i+1} - x_i} \end{pmatrix} \\ &= \frac{1}{2}\frac{1}{\Delta x_i} \begin{pmatrix} -\dfrac{2}{c}\sqrt{\frac{1}{2}\, g\, \frac{h_i+h_{i+1}}{2}}\,(H(x_{i+1}) - H(x_i)) \\ g\, \dfrac{h_i+h_{i+1}}{2}\,(H(x_{i+1}) - H(x_i)) \end{pmatrix} \end{aligned} \tag{30}$$

Then, from Equation (17) and the previous expressions we obtain

$$\mathcal{G}(x_{i-1},x_i,x_{i+1},\bar{W}_{i-1},\bar{W}_i,\bar{W}_{i+1})$$
$$=\frac{1}{2}\frac{1}{\Delta x}\begin{pmatrix}\frac{1}{c}\sqrt{g(h_i+h_{i-1})}(H(x_i)-H(x_{i-1}))-\frac{1}{c}\sqrt{g(h_i+h_{i+1})}(H(x_{i+1})-H(x_i))\\ \\ g\frac{h_{i-1}+h_i}{2}(H(x_i)-H(x_{i-1}))+g\frac{h_i+h_{i+1}}{2}(H(x_{i+1})-H(x_i))\end{pmatrix}$$

When $h\equiv H$, this expression coincides with (28) if $c_2=2$. Thus, in this case we obtain the enhanced consistency condition.

Let us next assume that the function $\tilde{G}$ is given by (26). Taking into account the preceding analysis, and considering that

$$H'\left(\frac{x+y}{2}\right)=\frac{H(y)-H(x)}{y-x}+\mathcal{O}(\Delta x^2)$$

we deduce that for this choice of $\tilde{G}$ our scheme satisfies the approximated consistency condition. □

3.1. Numerical tests

In general, testing numerical solvers for shallow water equations faces the problem that there are no known (up to the knowledge of the authors) analytic time-dependent solutions. Thus indirect testing is needed. This testing may focus either on qualitative properties of the flow, comparison with well-known numerical solvers, comparison with experimental measurements or with approximated analytical solution. Here, we shall test our scheme by a quantitative and a qualitative test, introduced in References [5, 6]. In Section 4, we shall also perform a test by comparison with experimental measurements in channels with constant widths, by taking into account friction effects, and another by comparison with Roe's Solver.

We shall consider tests corresponding to the depth profile function $H(x)$ defined as

$$H(x)=50.5-\frac{40x}{L}+10\,\sin\left(\pi\left(\frac{4x}{L}+\frac{1}{2}\right)\right) \tag{31}$$

and to initial and boundary conditions of the form

$$\begin{aligned}&h(x,0)=H(x),\quad q(x,0)=0\\&h(0,t)=\varphi(t)+H(0),\quad q(L,t)=\psi(t)\end{aligned} \tag{32}$$

where the functions $\varphi(t)$ and $\psi(t)$ are specified for each particular test.

Notice that the actual initial conditions that we consider correspond to flow at rest. Also, we shall consider in this section subcritical flows, so we set a Dirichlet boundary condition at $x=0$ and another at $x=L$. In the two tests below we impose h in $x=0$ and q in $x=L$. However, for the numerical method we need four boundary conditions. Therefore, we also impose $\partial_n q=0$ in $x=0$ and $\partial_n h=0$ in $x=L$. These are not physical boundary conditions, so they can produce some inaccuracies.

The numerical results that we present correspond to the numerical source term given by (25), for which the enhanced consistency condition is achieved, by Theorem 1.

Test 1 (*Quantitative test*): Let us, respectively, denote by $\mathscr{F}$ the *Froude* number and by $\mathscr{U}$ a non-dimensional velocity of the flow

$$\mathscr{F} = \frac{U}{\sqrt{gH^*}}, \quad \mathscr{U} = \frac{L}{TU}$$

where H^* and U denote the characteristic depth and velocity of the flow. In Reference [5] an asymptotic analysis of shallow water equations (1) with respect to the parameter $\mathscr{F}$ when $\mathscr{U} \approx 1$ is performed. Concretely, it is formally proved that, in the limit

$$\mathscr{F} \to 0$$

the solution (h, q) of (1) with initial conditions (32) approaches the pair of functions

$$\hat{h}(x,t) = \varphi(t) + H(x) \tag{33}$$

$$\hat{q}(x,t) = -\varphi'(t)(x - L) + \psi(t) \tag{34}$$

Notice that for this solution, the water surface level is constant at each fixed time.

In our test, we have taken

$$\varphi(t) = 4 + 4 \sin\left(\pi\left(\frac{4t}{86\,400} + \frac{1}{2}\right)\right)$$

$$\psi(t) = 0$$

We have considered space and time intervals $[0, L] = [0, 14\,000]$ (in m) and $[0, T] = [0, 10\,800]$ (in s). This approximately corresponds to the parameters $\mathscr{F} = 0.06$ (strongly subcritical flow). The time $t = 10\,800$ (3 h) corresponds to half downward tide.

The space interval $[0, L]$ has been divided into 50 subintervals. We have iterated in time our scheme with 0.8 for the CFL condition.

We present our results for time $t = T = 10\,800$. In Figure 1 we compare the computed velocity $u = q/h$ versus the analytical velocity $\hat{u} = \hat{q}/\hat{h}$. We observe a very good accuracy of our computation. Also, in Figure 2 we present the free surface, which in fact is very accurately constant along the interval $[0, L]$.

Test 2 (*Qualitative test*): In this test, we analyse the accuracy of our scheme in the computation of the speed of waves.

We consider a long domain, $L = 648\,000$ m (648 km), and analyse the penetration in the channel of a rising tide of 4 m of amplitude. This correspond to the initial conditions given by

$$\varphi(t) = 4 + 4 \cos\left(\pi\left(\frac{4t}{86\,400} + \frac{1}{2}\right)\right)$$

$$\psi(t) = 0$$

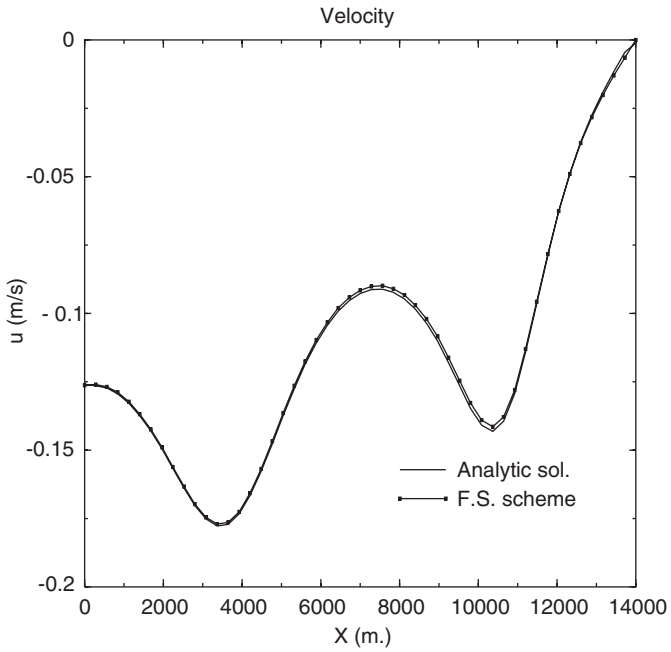

Figure 1. Test 1: velocity at $t = 10\,800$.

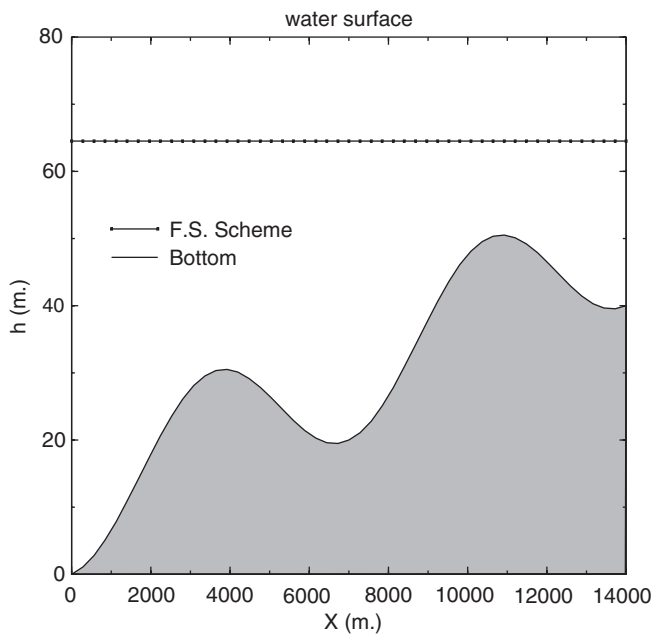

Figure 2. Test 1: free surface at $t = 10\,800$.

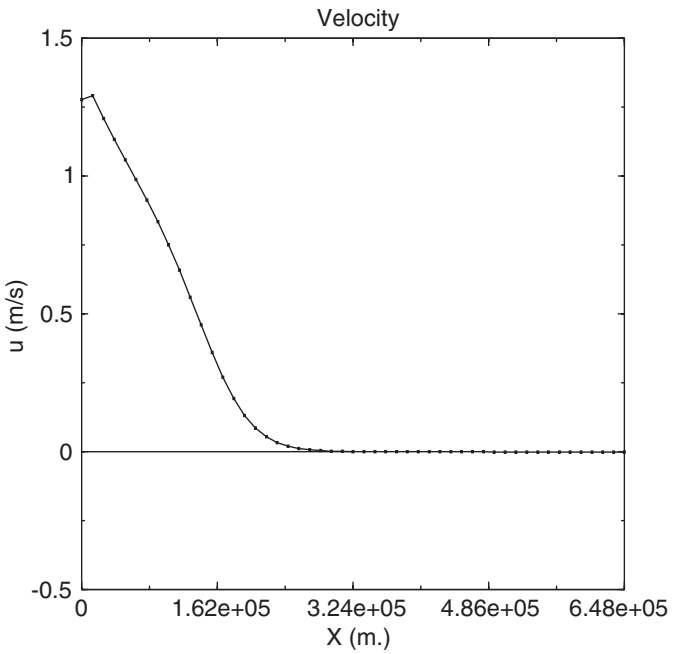

Figure 3. Test 2: velocity at $t = 10\,800$.

As the speed of small waves is

$$c = \sqrt{gh}$$

then in 10 800 s (3 h) the rising tide cannot penetrate in the channel more that about 216 km. We are still dealing with subcritical flow, as this test corresponds to a Froude number $\mathscr{F} \simeq 0.6$. However, this large value of $\mathscr{F}$ does not fit in the framework of the previous case.

We have taken $\Delta x = L/50 = 12\,960$ and CFL equal to 0.8, so we have iterated 2160 time steps up to half rising tide time, $t = 10\,800$ s.

In Figure 3, we observe that the velocity computed by our scheme effectively is zero for approximately $x \geqslant 216\,000$, as we expected. Also, the free surface (Figure 4) remains constant from this point.

Globally, the performance of our scheme for these analytical tests is quite close to that of the extension of Roe's scheme reported in Reference [5].

4. VARIABLE WIDTH AND FRICTION EFFECTS

In this section, we extend our solver to the case when the source terms take into account variable width of the channel and friction effects, in addition to variable depth. We prove

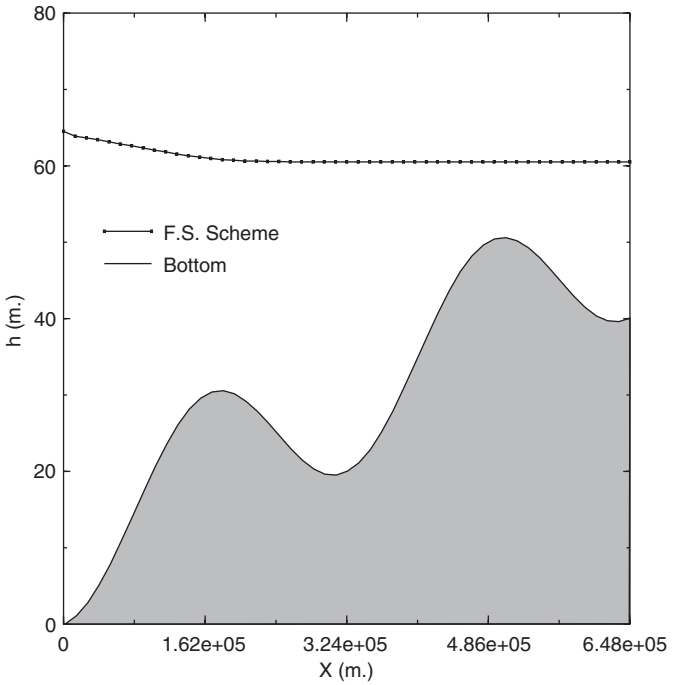

Figure 4. Test 2: free surface at $t = 10\,800$.

that a enhanced consistency condition is still satisfied, and perform some numerical tests that exhibit the performances of our solver for these source terms.

We shall consider a source term for Equation (1) with the following structure:

$$G(x, W) = G_1(x, W) + G_2(x, W) + G_3(x, W) \tag{35}$$

Here G_1 takes into account variable depth as in the preceding section

$$G_1(x, (h, q)) = \begin{pmatrix} 0 \\ ghH'(x) \end{pmatrix} \tag{36}$$

Also, G_2 represents the action of variable width on the flow

$$G_2(x, (h, q)) = \begin{pmatrix} -q \dfrac{b'(x)}{b(x)} \\ -\dfrac{q^2}{h} \dfrac{b'(x)}{b(x)} \end{pmatrix} \tag{37}$$

where $b(x)$ is the width of the channel at position x. We shall assume that $b(x)$ is a positive and twice continuously differentiable function on $[0, L]$.

G_3 represents the bottom and sidewalls friction effects

$$G_3(x,(h,q)) = \begin{pmatrix} 0 \\ -gq\left|\frac{q}{h}\right|\left(\frac{M_b^{3/2}}{h} + \frac{2M_w^{3/2}}{b(x)}\right)^{4/3} \end{pmatrix} \tag{38}$$

where M_b and M_w, respectively, are the bottom and wall Manning coefficients (Cf. Reference [12]).

Corresponding to the structure of the source term (35), we shall consider a numerical source term of the form

$$\mathcal{G}(x_{i-1},x_i,x_{i+1},W_{i-1},W_i,W_{i+1}) = (\mathcal{G}_1+\mathcal{G}_2+\mathcal{G}_3)(x_{i-1},x_i,x_{i+1},W_{i-1},W_i,W_{i+1}) \tag{39}$$

where $\mathcal{G}_1$, is still given by (17)–(22). Also, to define $\mathcal{G}_2$ and $\mathcal{G}_3$ we use the same idea as for the construction of $\mathcal{G}_1$ but, in addition, we shall differently treat each equation (similarly to the construction of the numerical flux ϕ). We concretely define the lth component of the $\mathcal{G}_j$, $j=1,2,3$, by

$$\begin{aligned}
[\mathcal{G}_j(x_{i-1},x_i,x_{i+1},W_{i-1},W_i,W_{i+1})]_l &= [\mathcal{G}_{j\mathrm{L}}]_l + [\mathcal{G}_{j\mathrm{R}}]_l \\
&= \left[\frac{(x_i-x_{i-1})/2}{\Delta x_i}(I+|B_j^l(W_{i-1},W_i)|(B_j^l)^{-1}(W_{i-1},W_i))\tilde{G}_j(x_{i-1},x_i,W_{i-1},W_i)\right]_l \\
&\quad + \left[\frac{(x_{i+1}-x_i)/2}{\Delta x_i}(I-|B_j^l(W_i,W_{i+1})|(B_j^l)^{-1}(W_i,W_{i+1}))\tilde{G}_j(x_i,x_{i+1},W_i,W_{i+1})\right]_l
\end{aligned} \tag{40}$$

Here, we define B_j^l for $j=1,2,3$, $l=1,2$ by

$$B_j^l(W_i,W_{i+1}) = \tilde{A}_j^l\left(\frac{W_i+W_{i+1}}{2}\right)$$

with

$$\tilde{A}_j^l = \tilde{X}_j^l \Lambda (\tilde{X}_j^l)^{-1}$$

where

$$\tilde{X}_j^l = \begin{pmatrix} 1 & 0 \\ 0 & c_{j,l} \end{pmatrix} X$$

for some real numbers $c_{j,l}$ to determine. Due to the analysis of the preceding section, we set $c_{1,1}=c_{1,2}=2$.

Also, the functions $\tilde{G}_j(x_i, x_{i+1}, W_i, W_{i+1})$ are approximations of $G_j(x_{i+1/2}, W_{i+1/2})$, $j = 1, 2, 3$. We define $\tilde{G}_1$ by (25), and set

$$\tilde{G}_2(x_i, x_{i+1}, W_i, W_{i+1}) = \begin{pmatrix} -(q_i + q_{i+1}) \dfrac{1}{x_{i+1} - x_i} \dfrac{b(x_{i+1}) - b(x_i)}{b(x_{i+1}) + b(x_i)} \\ -\dfrac{(q_i + q_{i+1})^2}{h_i + h_{i+1}} \dfrac{1}{x_{i+1} - x_i} \dfrac{b(x_{i+1}) - b(x_i)}{b(x_{i+1}) + b(x_i)} \end{pmatrix} \tag{41}$$

and

$$\tilde{G}_3(x_i, x_{i+1}, W_i, W_{i+1}) = \begin{pmatrix} 0 \\ -g \dfrac{q_i + q_{i+1}}{2} \left| \dfrac{q_i + q_{i+1}}{h_i + h_{i+1}} \right| \left(\dfrac{2M_b^{3/2}}{h_i + h_{i+1}} + \dfrac{4M_w^{3/2}}{b(x_i) + b(x_{i+1})} \right)^{4/3} \end{pmatrix} \tag{42}$$

We recall that $\tilde{G}_3(x_i, x_{i+1}, W_i, W_{i+1})$ is an approximation of

$$\frac{1}{x_{i+1} - x_i} \int_{x_i}^{x_{i+1}} G_3(x, W)\, \mathrm{d}x$$

Definition (42) uses a modification of the mid point rule to approximate this integral. It is a second-order accurate cuadrature rule.

The enhanced consistency condition for this case cannot be referred to the solution $h = H$, $q = 0$ as for these functions $G_2 = 0$ and $G_3 = 0$. We instead use the stationary solution of Equations (1) with source term (35) reported in Reference [13],

$$\bar{W}(x) = \begin{pmatrix} \bar{h} \\ \dfrac{\bar{h}k}{b(x)} \end{pmatrix} \tag{43}$$

where $\bar{h}$ and k are arbitrary constants. We shall assume $\bar{h} > 0$, as it represents the height of a water column. The depth function H is not arbitrary here, but instead it is determined in order to have (43) solution of (1) with source term (35),

$$H(x) = H(0) + \frac{k^2}{2g} \left(\frac{1}{b^2(x)} - \frac{1}{b^2(0)} \right) + k|k| \int_0^x \frac{1}{b^2(s)} \left(\frac{M_b^{3/2}}{h} + \frac{2M_w^{3/2}}{b(s)} \right)^{4/3} \mathrm{d}s \tag{44}$$

We may now state our main result.

Theorem 2

The flux-splitting scheme defined by (4), (15), (17)–(22) with $c_{1,1} = c_{1,2} = 2$, and (40)–(42) verifies the following:

1. In the absence of friction terms, the scheme satisfies the exact Bermúdez–Vázquez condition with respect to the stationary solution (43) if $c_{2,1} = \frac{2}{3}$ and $c_{2,2} = 2$.

2. For general source terms given by (35), the scheme satisfies the approximated Bermúdez–Vázquez condition with respect to the stationary solution (43) if in addition $c_{3,1} = c_{3,2} = 2$.

Proof

We start by calculating the expressions for the numerical flux and numerical source terms for general source terms. This will be next applied to each particular case.

Again, we shall denote $\bar{W}(x)$ by $\bar{W}$ and $\bar{W}(x_i)$ by $\bar{W}_i$.

To derive the expression for the numerical flux, let us recall that the eigenvalues of $A(\bar{W})$ are

$$\lambda_1 = \frac{k}{b} + \sqrt{\frac{g\bar{h}}{2}}, \qquad \lambda_2 = \frac{k}{b} - \sqrt{\frac{g\bar{h}}{2}}$$

The actual expressions for $A^+(\bar{W})$ and $A^-(\bar{W})$ will depend on which λ_i are positive. We shall treat here the case $\lambda_1 > 0$, $\lambda_2 < 0$. The other possible cases ($\lambda_1 > 0$, $\lambda_2 > 0$ and $\lambda_1 < 0$, $\lambda_2 < 0$) are straightforward.

When $\lambda_1 > 0$, $\lambda_2 < 0$, using (7), (8) and (10) we deduce

$$A^+(\bar{W}) = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} \lambda_1\lambda_2 & -\lambda_1 \\ \lambda_2(\lambda_1)^2 & -(\lambda_1)^2 \end{pmatrix}$$

$$A^-(\bar{W}) = \frac{1}{\lambda_2 - \lambda_1} \begin{pmatrix} -\lambda_1\lambda_2 & \lambda_2 \\ -\lambda_1(\lambda_2)^2 & (\lambda_2)^2 \end{pmatrix}$$

Observe that $\lambda_1 \neq \lambda_2$ as $\bar{h} > 0$. From the preceding expressions

$$(a_1)^+(\bar{W}) = \frac{-1}{\sqrt{2g\bar{h}}} \left(\frac{k^2}{b^2} - \frac{g\bar{h}}{2}, -\frac{k}{b} - \sqrt{\frac{g\bar{h}}{2}} \right)$$

$$(a_2)^+(\bar{W}) = \frac{-1}{\sqrt{2g\bar{h}}} \left(\left(\frac{k}{b} - \frac{g\bar{h}}{2} \right) \left(\frac{k}{b} + \frac{g\bar{h}}{2} \right)^2, -\left(\frac{k}{b} + \sqrt{\frac{g\bar{h}}{2}} \right)^2 \right)$$

$$(a_1)^-(\bar{W}) = \frac{-1}{\sqrt{2g\bar{h}}} \left(-\frac{k^2}{b^2} + \frac{g\bar{h}}{2}, \frac{k}{b} - \sqrt{\frac{g\bar{h}}{2}} \right)$$

$$(a_2)^-(\bar{W}) = \frac{-1}{\sqrt{2g\bar{h}}} \left(\left(-\frac{k}{b} - \frac{g\bar{h}}{2} \right) \left(\frac{k}{b} - \frac{g\bar{h}}{2} \right)^2, \left(\frac{k}{b} - \sqrt{\frac{g\bar{h}}{2}} \right)^2 \right)$$

Consequently, the numerical flux is given by

$$\phi(\bar{W}_i;\bar{W}_{i+1}) = \left(\frac{(a_1)^+\left(\bar{h}, \frac{k\bar{h}}{2}\left(\frac{1}{b_i}+\frac{1}{b_{i+1}}\right)\right)}{(a_2)^+\left(\bar{h}, \frac{k\bar{h}}{b_i}\right)} \right)\bar{W}_i + \left(\frac{(a_1)^-\left(\bar{h}, \frac{k\bar{h}}{2}\left(\frac{1}{b_i}+\frac{1}{b_{i+1}}\right)\right)}{(a_2)^-\left(\bar{h}, \frac{k\bar{h}}{b_{i+1}}\right)} \right)\bar{W}_{i+1}$$

Then,

$$\phi(\bar{W}_i,\bar{W}_{i+1}) - \phi(\bar{W}_{i-1},\bar{W}_i) = \begin{pmatrix} \frac{k\bar{h}}{2}\left(\frac{1}{b_{i+1}}-\frac{1}{b_{i-1}}\right) - \frac{k^2}{2g}\sqrt{\frac{g\bar{h}}{2}}\left(\frac{1}{b_{i+1}^2}+\frac{1}{b_{i-1}^2}-\frac{2}{b_i^2}\right) \\ \frac{k^2\bar{h}}{2}\left(\frac{1}{b_{i+1}^2}-\frac{1}{b_{i-1}^2}\right) - k\bar{h}\sqrt{\frac{g\bar{h}}{2}}\left(\frac{1}{b_{i+1}}+\frac{1}{b_{i-1}}-\frac{2}{b_i}\right) \end{pmatrix} \tag{45}$$

We next calculate the numerical source term. The functions $\tilde{G}_1$, $\tilde{G}_2$ and $\tilde{G}_3$, respectively, defined in (25), (41) and (42), are given by

$$(x_{i+1}-x_i)\tilde{G}_1(x_i,x_{i+1},\bar{W}_i,\bar{W}_{i+1}) = \begin{pmatrix} 0 \\ \bar{h}\frac{k^2}{2}\left(\frac{1}{b^2(x_{i+1})}-\frac{1}{b^2(x_i)}\right) + g\bar{h}k|k|\int_{x_i}^{x_{i+1}} \frac{1}{b^2(s)}\left(\frac{M_{\mathrm{b}}^{3/2}}{\bar{h}}+\frac{2M_{\mathrm{w}}^{3/2}}{b(s)}\right)^{4/3} \mathrm{d}s \end{pmatrix}$$

$$(x_{i+1}-x_i)\,\tilde{G}_2(x_i,x_{i+1},\bar{W}_i,\bar{W}_{i+1}) = \begin{pmatrix} k\bar{h}\left(\frac{1}{b_{i+1}}-\frac{1}{b_i}\right) \\ \frac{k^2\bar{h}}{2}\left(\frac{1}{b_{i+1}^2}-\frac{1}{b_i^2}\right) \end{pmatrix}$$

$$\tilde{G}_3(x_i,x_{i+1},\bar{W}_i,\bar{W}_{i+1}) = \begin{pmatrix} 0 \\ -g\bar{h}k|k|\left(\frac{1}{2}\left(\frac{1}{b_i}+\frac{1}{b_{i+1}}\right)\right)^2\left(\frac{M_{\mathrm{b}}^{3/2}}{\bar{h}}+\frac{4M_{\mathrm{w}}^{3/2}}{b_i+b_{i+1}}\right)^{4/3} \end{pmatrix}$$

Also,

$$I + |B_j^l(\bar{W}_i, \bar{W}_{i+1})|(B_j^l)^{-1}(\bar{W}_i, \bar{W}_{i+1})$$

$$= -\sqrt{\frac{2}{g\bar{h}}} \begin{pmatrix} \frac{k}{2}\left(\frac{1}{b_{i-1}} + \frac{1}{b_i}\right) - \sqrt{\frac{g\bar{h}}{2}} & \frac{-1}{c_{j,l}} \\ c_{j,l}\left(\frac{k^2}{4}\left(\frac{1}{b_{i-1}} + \frac{1}{b_i}\right)^2 - \frac{g\bar{h}}{2}\right) & \frac{-k}{2}\left(\frac{1}{b_{i-1}} + \frac{1}{b_i}\right) - \sqrt{\frac{g\bar{h}}{2}} \end{pmatrix}$$

$$I - |B_j^l(\bar{W}_i, \bar{W}_{i+1})|(B_j^l)^{-1}(\bar{W}_i, \bar{W}_{i+1})$$

$$= -\sqrt{\frac{2}{g\bar{h}}} \begin{pmatrix} \frac{-k}{2}\left(\frac{1}{b_{i+1}} + \frac{1}{b_i}\right) - \sqrt{\frac{g\bar{h}}{2}} & \frac{1}{c_{j,l}} \\ -c_{j,l}\left(\frac{k^2}{4}\left(\frac{1}{b_{i+1}} + \frac{1}{b_i}\right)^2 - \frac{g\bar{h}}{2}\right) & \frac{k}{2}\left(\frac{1}{b_{i+1}} + \frac{1}{b_i}\right) - \sqrt{\frac{g\bar{h}}{2}} \end{pmatrix}$$

We next consider separately the cases with and without friction effects.

Step 1: If we do not take into account friction effects ($M_b = M_w = 0$), then, by (40),

$$\Delta x_i \mathscr{G}(x_{i-1}, x_i, x_{i+1}, \bar{W}_{i-1}, \bar{W}_i, \bar{W}_{i-1}) = \Delta x_i(\mathscr{G}_1 + \mathscr{G}_2)(x_{i-1}, x_i, x_{i+1}, \bar{W}_{i-1}, \bar{W}_i, \bar{W}_{i-1})$$

$$= \begin{pmatrix} \frac{k\bar{h}}{2}\left(\frac{1}{b_{i+1}} - \frac{1}{b_{i-1}}\right) - \frac{k^2}{g}\sqrt{\frac{g\bar{h}}{2}}\left(\frac{1}{2}\left(\frac{1}{c_{2,1}} - 1\right) + \frac{1}{4}\right)\left(\frac{1}{b_{i-1}^2} + \frac{1}{b_{i+1}^2} - \frac{2}{b_i^2}\right) \\ \frac{k^2\bar{h}}{2}\left(\frac{1}{b_{i+1}^2} - \frac{1}{b_{i-1}^2}\right) - \frac{c_{2,2}}{2}k\bar{h}\sqrt{\frac{g\bar{h}}{2}}\left(\frac{1}{b_{i-1}} + \frac{1}{b_{i+1}} - \frac{2}{b_i}\right) + \sqrt{\frac{g\bar{h}}{2}}\frac{k^3}{4g}(c_{2,2} - 2) \\ \left(\left(\frac{1}{b_{i+1}} + \frac{1}{b_i}\right)^2\left(\frac{1}{b_{i+1}} - \frac{1}{b_i}\right) - \left(\frac{1}{b_i} + \frac{1}{b_{i-1}}\right)^2\left(\frac{1}{b_i} - \frac{1}{b_{i-1}}\right)\right) \end{pmatrix}$$

This expression must coincide with (45) in order the enhanced consistency condition to hold. This requires the following three equations (with two unknowns) to hold

$$\frac{1}{2}\left(\frac{1}{c_{2,1}} - 1\right) + \frac{1}{4} = \frac{1}{2}$$

$$\frac{c_{2,2}}{2} = 1$$

$$c_{2,2} - 2 = 0$$

The first equation holds with $c_{2,1} = \frac{2}{3}$, while the second and third ones hold with $c_{2,2} = 2$.

Step 2: The difference of numerical fluxes is balanced by the numerical source terms G_1 and G_2 of step 1. When we include friction effects, due to the definition of H (44), a new summand $\bar{\mathscr{G}}_1$ appear in $\mathscr{G}_1$. Then, it suffices to prove that $\bar{\mathscr{G}}_1 + \mathscr{G}_3 = \mathscr{O}(\Delta x^2)$.

We compute the right and left components of $\bar{\mathscr{G}}_1 = \bar{\mathscr{G}}_{1,\mathrm{L}} + \bar{\mathscr{G}}_{1,\mathrm{R}}$ and $\mathscr{G}_3 = \mathscr{G}_{3,\mathrm{L}} + \mathscr{G}_{3,\mathrm{R}}$,

$$[\bar{\mathscr{G}}_{1,\mathrm{R}} + \mathscr{G}_{3,\mathrm{R}}](x_i, x_{i+1}, \bar{W}_i, \bar{W}_{i+1}) = \frac{1}{\Delta x_i} \begin{pmatrix} g\bar{h}k|k| \dfrac{1}{\sqrt{2g\bar{h}}} \left(\dfrac{1}{c_{3,1}} M_i - \dfrac{1}{c_{1,1}} N_i \right) \\ g\bar{h}k|k| \left[\dfrac{k}{2} \left(\dfrac{1}{b_i} + \dfrac{1}{b_{i+1}} \right) - \sqrt{\dfrac{g\bar{h}}{2}} \right] (M_i - N_i) \end{pmatrix}$$

where

$$M_i = (x_{i+1} - x_i) \left[\frac{1}{2} \left(\frac{1}{b_i} + \frac{1}{b_{i+1}} \right) \right]^2 \left(\frac{M_{\mathrm{b}}^{3/2}}{\bar{h}} + \frac{4M_{\mathrm{w}}^{3/2}}{b_i + b_{i+1}} \right)^{4/3}$$

$$N_i = \int_{x_i}^{x_{i+1}} \frac{1}{b^2(s)} \left(\frac{M_{\mathrm{b}}^{3/2}}{\bar{h}} + \frac{2M_{\mathrm{w}}^{3/2}}{b(s)} \right)^{4/3} \mathrm{d}s$$

M_i is the result of applying a modification of the midpoint numerical integration rule to the integral appearing in N_i. As the integrand is a C^2 function and this formula is second-order accurate, taking $c_{3,1} = c_{1,1} = 2$ we have $\bar{\mathscr{G}}_{1,\mathrm{R}} + \mathscr{G}_{3,\mathrm{R}} = O(\Delta x^2)$. Taking $c_{3,2} = 2$, with a similar deduction, we prove that $\bar{\mathscr{G}}_{1,\mathrm{L}} + \mathscr{G}_{3,\mathrm{L}} = O(\Delta x^2)$. Then, the approximated consistency condition is verified. □

Remark 1

In the last case, using a more accurate quadrature formula for the integrals defining the B_i coefficients, we may obtain an approximated condition of order $O(\Delta x^p)$, for $p > 2$ only limited by the regularity of the function $b(x)$.

4.1. Numerical tests

We present the results of two tests for the flux-splitting scheme with complex source terms. The first one compares the solution provided by the scheme with that of an analytical approximation for small Froude numbers, similar to the test of the preceding section. The second one compares our scheme with experimental measurements of a dam breaking experiment. We include a comparison with the results provided by Roe's solver.

Test 3 (*Analytical test*): In Reference [13] it is also reported a limit solution of shallow water equations (1) with source terms (35), for small Froude numbers and 'short' domains. Concretely, when $\mathscr{F} \to 0$ and $\mathscr{U} \simeq 1$, the solution (h, q) of (1), (35) formally approaches the pair of functions

$$\hat{h}(x,t) = \varphi(t) + H(t), \quad \hat{q}(x,t) = \psi(t) + \frac{\varphi'(t)}{b(x)} \int_x^L b(s)\,\mathrm{d}s$$

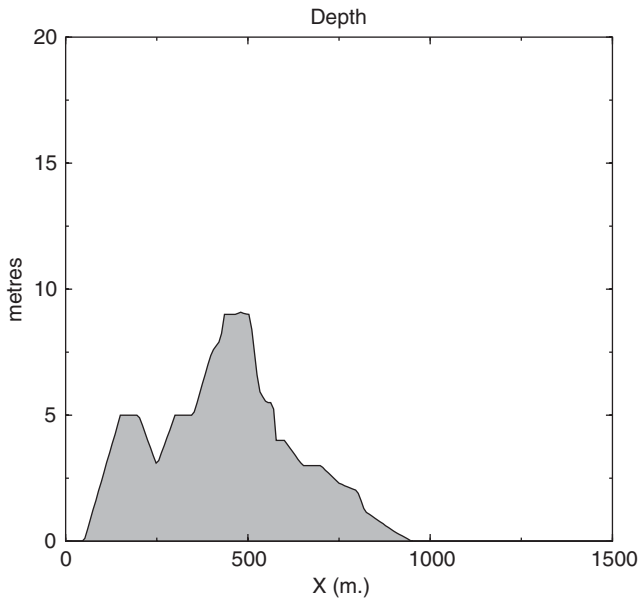

Figure 5. Test 3: depth function.

Observe that also for this solution the surface water level is horizontal at each fixed time.

In our test we have taken the initial and boundary conditions as in (32),

$$\begin{aligned} &h(x,0)=H(x), \quad q(x,0)=0 \\ &h(0,t)=\varphi(t)+H(0), \quad q(L,t)=\psi(t) \end{aligned} \tag{46}$$

with

$$\varphi(t)=4+4\sin\left[\pi\left(\frac{4t}{86\,400}-\frac{1}{2}\right)\right]; \quad \psi(t)=0$$

The boundary data correspond to a rising tide with amplitude of 8 m at $x=0$ and to a zero discharge at $x=L$. Also, we have taken $L=1500$ (in m) and $T=10\,800$ (in s, half rising tide). This corresponds to the parameter $\mathscr{F}\simeq 0.1$, so we may consider that we are in the asymptotic regime $\mathscr{F}\to 0$.

We have taken $\Delta x=7.5$ and $\Delta t=0.4$. The Manning coefficients for bottom and sidewalls taken are

$$M_{\mathrm{b}}=0.1, \qquad M_{\mathrm{w}}=0.18$$

Also, we have taken the profile bottom and width functions proposed in Reference [13], these are, respectively, represented in Figures 5 and 6.

As the flow for this test is subcritical, we impose the same boundary conditions as in Tests 1 and 2.

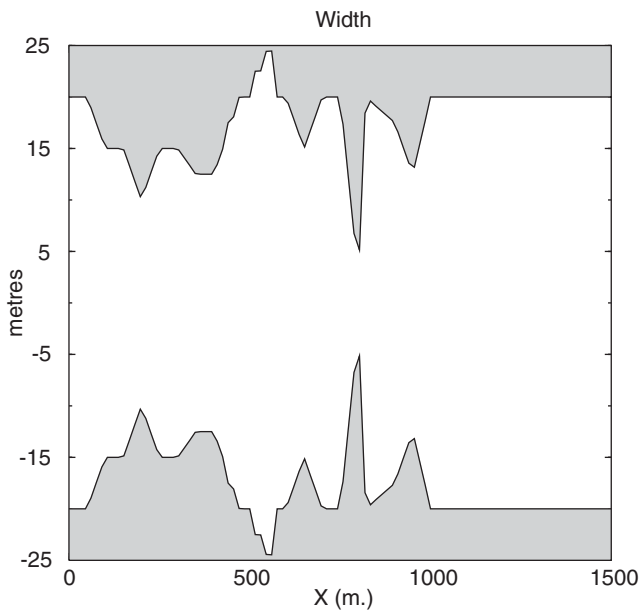

Figure 6. Test 3: width function.

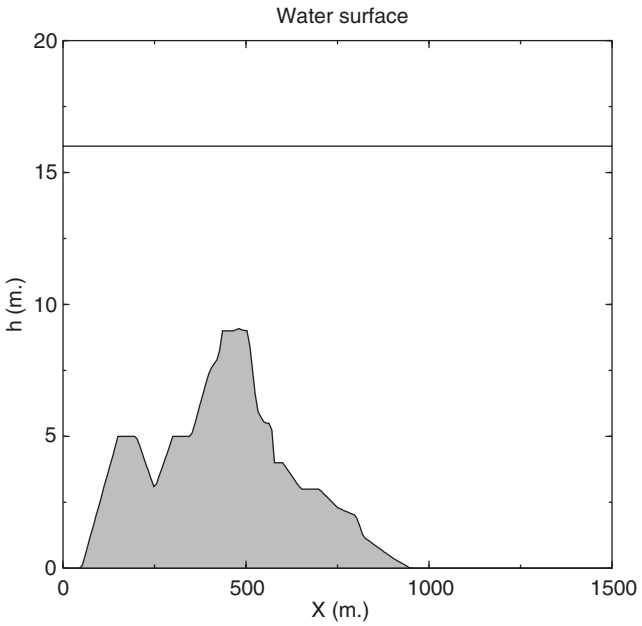

Figure 7. Test 3: computed water surface at $t = 10\,800$.

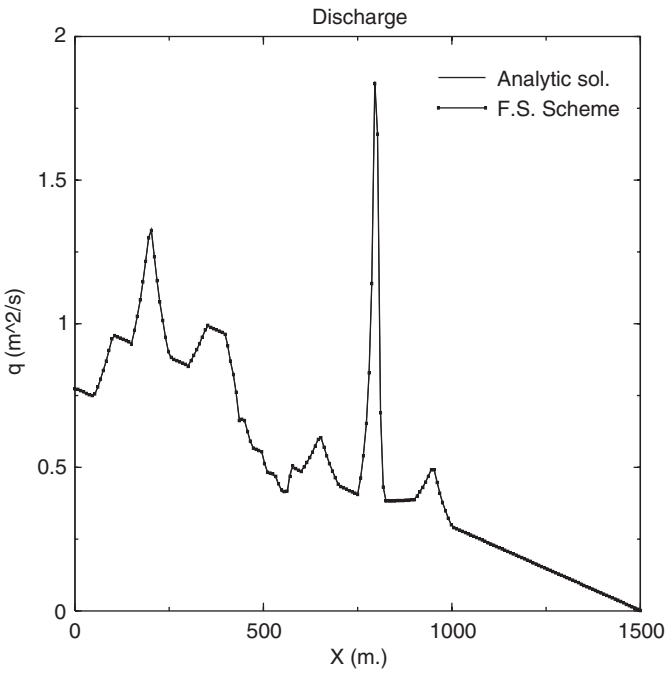

Figure 8. Test 3: computer versus analytical discharge at $t = 10\,800$.

We, respectively, present in Figures 7 and 8 the computed free surface and the discharge, compared with the analytical solution, at time $t = T = 10\,800$. We observe that the free surface is effectively horizontal. The maximum error between computed and analytical solution at grid nodes is 2×10^{-3}.

Test 4 (*Dam break experiment*): Our last test consists in comparing the results of our solver with the measurements of a dam break experiment, performed in the Laboratoire de Recherches Hydrauliques of the Université Libre de Bruxelles, under the direction of Prof. J. M. Hiver. These measurements have been used in Reference [12] to test the extension of Roe's solver developed by Vázquez Cendón in Reference [6].

In this test h can be equal to zero (dry zones), so a special discretization of the source term G_1 should be used, in such a way that no artificial pressure is introduced. We use a technical trick proposed in Reference [12]. Essentially, this trick consists in modifying the bottom slope in cells with dry/wet zones, in order to respect the mass conservation.

Certainly, comparison with experimental measurements is the ultimate test for a numerical model, as it yields its true accuracy with respect to the physical behaviour of the flow. Nevertheless, in our case it must be considered as a relative test, as the solver is only intended to approximate the solutions of the continuous shallow water equations model (1). This model does not contain internal diffusion effects, neither turbulence modelling. Moreover, its derivation is based upon the hydrostatic pressure hypothesis. Therefore, the momentum conservation

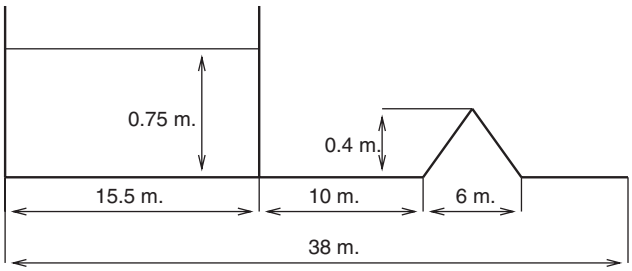

Figure 9. Test 4: channel profile. Free outflow.

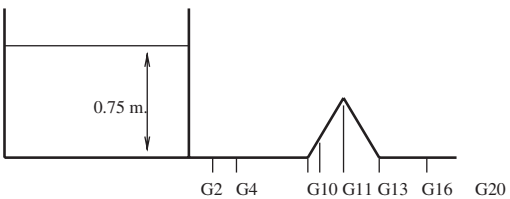

Figure 10. Test 4: measurement points

equation of model (1) would be far from representing the true momentum balance in zones where turbulent effects take place, or the flow evolves in horizontal and vertical scales of similar sizes.

The testing of accuracy should be conditioned by these facts. We may, however, test the accuracy of the computations of the speed of surface travelling waves, which is not highly affected by diffusion effects. We may also test the accuracy of our solver by comparison with other solvers with good stability and accuracy properties.

The experiment is as follows. A reservoir of 15.5 meters length is filled with water up to a height of 0.75 m. A floodgate separates the reservoir from a straight channel of 22.5 m length, with a triangular obstacle of 0.4 m height and 6 m length (See Figure 9). The width is constant 1.75 m. Three kinds of outflow boundary are considered: Free exit (Test 4.1, Figure 10), short vertical wall (Test 4.2, Figure 13), large vertical wall (Test 4.3, Figure 16). The experimental data available are the free surface position along the time interval $[0, 40]$ (s), at 20 points, represented in Figure 10. The Manning coefficients for bottom and sidewalls are $M_b = 0.0125$ and $M_w = 0.011$.

In our experiments we have taken 152 points along the channel, corresponding to $\Delta x =$ 25 cm. We have adapted the second-order four-step explicit Runge–Kutta scheme used in Reference [14], for the time discretization.

To solve the Cauchy problem

$$y'(t) = f(t, y(t))$$

$$y(t_0) = y_0$$

this scheme updates y^n to y^{n+1} in four intermediates steps:

$$y^{n,0} = y^n \quad t_{n,0} = t_n$$
$$y^{n,k} = y^n + \alpha_k \Delta t^n f(t_{n,k-1}, y^{n,k-1})$$
$$t_{n,k} = t_n + \alpha_k \Delta t^n \quad k = 1, 2, 3, 4$$
$$y^{n+1} = y^{n,4} \quad t_{n+1} = t_{n,4}$$

with

$$\alpha_1 = 0.11, \quad \alpha_2 = 0.2766, \quad \alpha_3 = 0.5, \quad \alpha_4 = 1$$

This provides a solver with an overall second-order accuracy. Thus, we may expect that the discrepancies between numerical results and experimental measurements are mainly due to the continuous model, rather than to the numerical solver. The increase of computational complexity involved for using this second-order solver is not high, as it provides stable solutions with CFL numbers up to 2.5, while the explicit Euler scheme is stable for CFL numbers up to 0.8.

At the solid wall (located at $x = 0$) we impose the physical boundary conditions $q = 0$ and also $\partial_n h = 0$.

We, respectively, represent in Figures 12, 15, 18 the water surface at $t = 20$s, and in Figures 11–16 and 17, 18 the time evolution of the computed free surface at points G10, G11, G13 and G20 for sub-tests 4.1–4.3 below. These points likely are the more meaningful, as the first three are situated along the obstacle (in particular, the top of the obstacle is situated at G13), and the fourth one is between the obstacle and the outflow boundary.

Test 4.1 (*Free outflow condition*): We have simulated this case by imposing the conditions

$$\partial_n h = 0, \quad \partial_n q = 0$$

at the outflow boundary ($x = 38$).

We may observe a good accuracy at all points considered. The speed of propagation of discontinuities is well computed. Even the overall pattern of the free surface evolution at the top of the obstacle (point G13) is well reproduced. In general, the computed heights are larger than the physical ones. We think that this is a consequence of the lack of energy dissipation mechanisms in our model.

Test 4.2 (*Short vertical wall at outflow boundary*): In this case the end section of the channel, between the obstacle and the outflow boundary, is assumed to be filled with water at rest at time $t = 0$. There exits an engineering law which models the outflow boundary condition,

$$q = C_w (h - h_v)^{2/3}$$

where h_v is the height of the outflow vertical wall (in our case, $h_v = 15$ cm) and C_w is an empirical constant (usually, $C_w = 1.7$). In our case, we have preferred to simulate the vertical wall as a part of the bottom function with an infinite gradient. Numerically, this gradient is

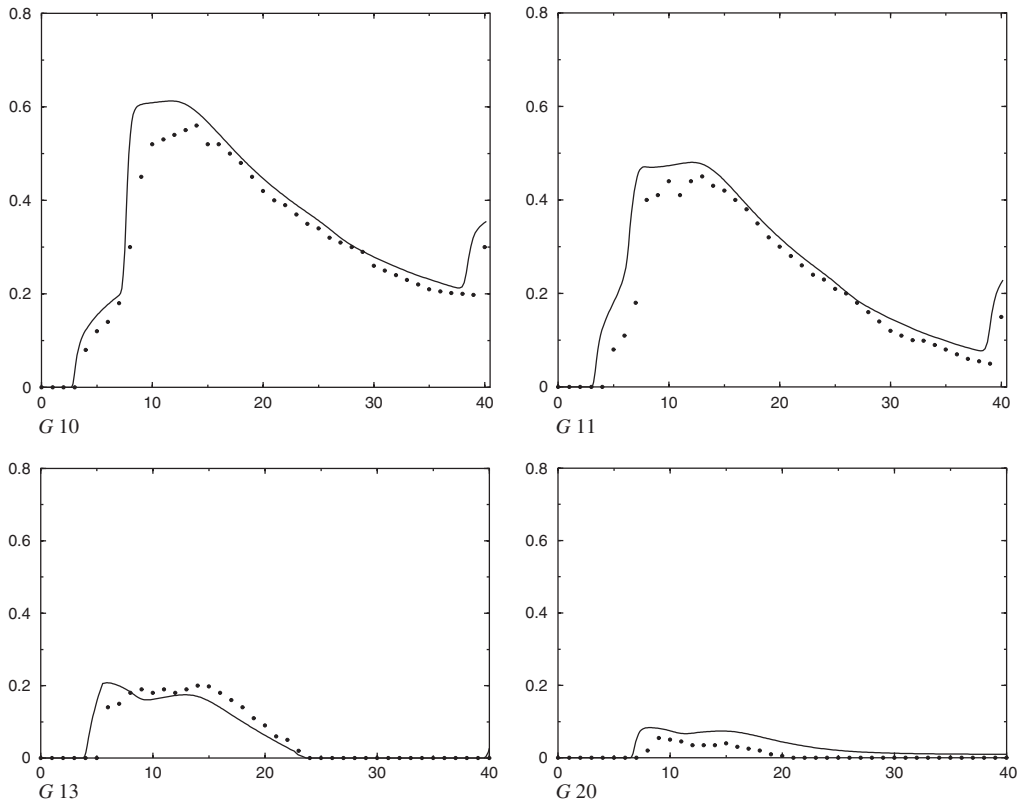

Figure 11. Test 4.1: free surface evolution at measurement points. Solid line: numerical results. Dotted line: experimental measurements.

taken as $h_v/\Delta x$. At the last out flow point we have set $\partial_n q = \partial_n h = 0$. This provides practically the same results.

In this case, we still recover a good accuracy for our computed free surface (Figure 14). The same comments as for Test 4.1 still apply here. Moreover, there exist small reflections due to the presence of the outflow wall. This is made apparent in the history at point G20, where we may observe the passage of two fronts, whose velocities are accurately simulated.

Test 4.3 (*Large vertical wall at outflow boundary*, *Comparison with Roe's solver*): In this case, we have set the boundary conditions $q = 0$, $\partial_n h = 0$ to simulate the effect of the outflow wall. The presence of the wall at the outflow boundary produces several reflections which travel back and forth along the channel and are in their turn reflected by the obstacle.

The overall pattern of the time behaviour of the free surface is again correctly represented by the numerical solution. The four fronts passing along the point G13 are recovered, so as their speed of propagation, and the zero depth in the time interval [29,33], approximately. In point G20 these characteristics also are correctly simulated, with a quite satisfactory accuracy.

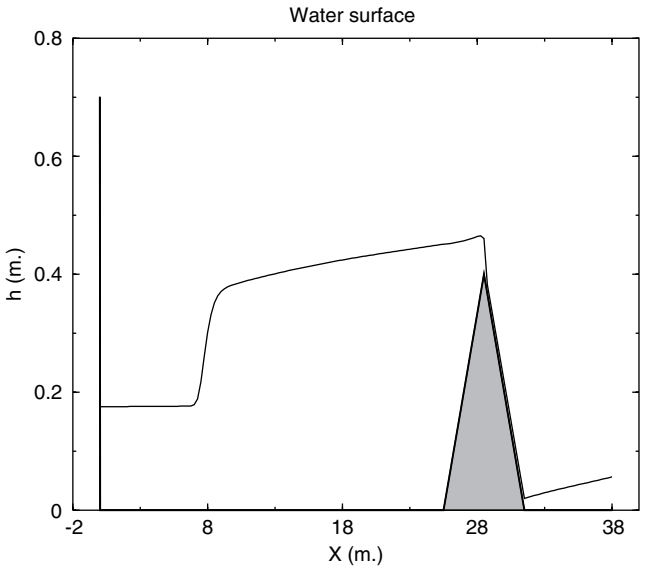

Figure 12. Test 4.1: computed water surface at $t = 20$ s.

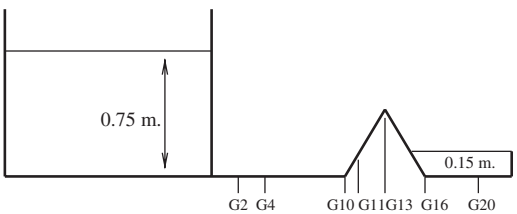

Figure 13. Test 4.2: channel profile. Short vertical wall at outflow boundary.

However, for the measurement points situated G10 and G13, there are relatively large deviations with respect to the experimental measurements. This probably occurs at first because there is no loss of fluid across an outflow boundary, and then the total energy of the system is conserved. Very likely, this produces high levels of turbulence in this flow, especially close to these points.

Moreover, the point G10 is situated in a corner where possibly the boundary layer detaches and a recirculation zone takes place. This would invalidate the hydrostatic pressure assumption in this zone. This could explain the larger discrepancies between experimental measurements and numerical results at point G10, with respect to the other points considered.

We also compare for this Test 4.3 the solution provided by our solver with that provided by the extension of Roe's solver introduced in Reference [13]. In general, both solvers provide very close results.The largest discrepancies between both solvers occur at point G13, which correspond to the top of the obstacle. This may be due to the high irregularity of the flow at this point.

We also stress that the computational cost of both solvers are quite similar.

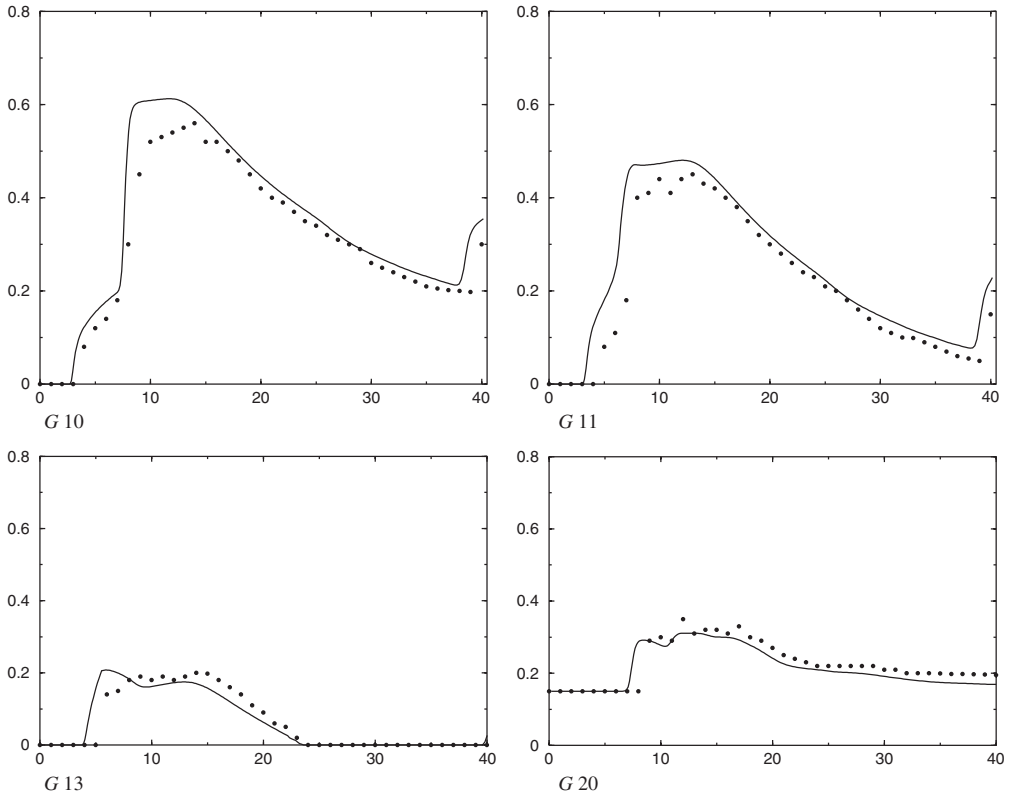

Figure 14. Test 4.2: free surface evolution at measurement points. Solid line: numerical results. Dotted line: experimental measurements.

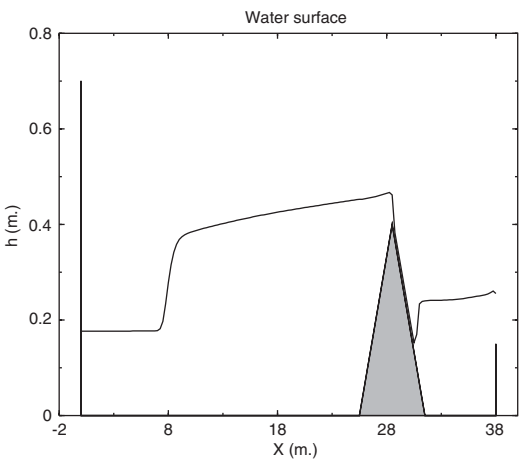

Figure 15. Test 4.2: computed water surface at $t = 20$ s.

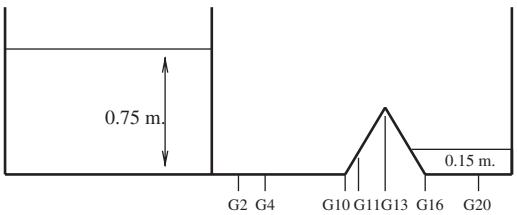

Figure 16. Test 4.3: channel profile. Large vertical wall at outflow boundary.

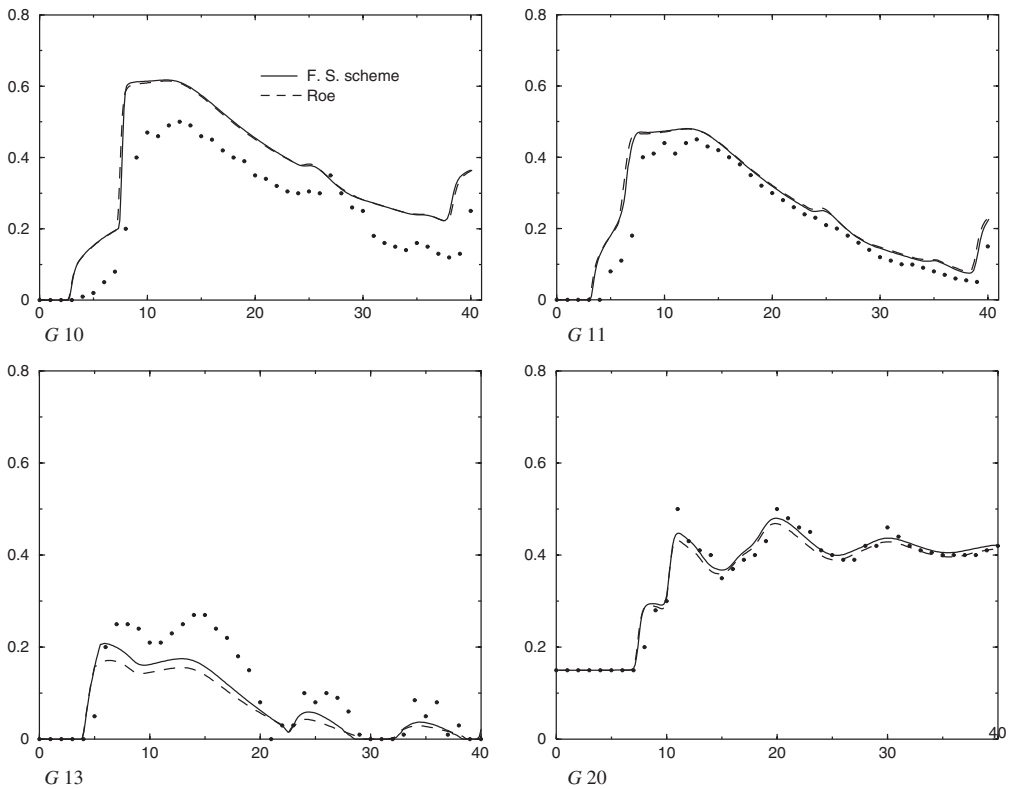

Figure 17. Test 4.3: free surface evolution at measurement points. Solid line: flux-splitting scheme. Dashed line: Roe's scheme. Dotted line: experimental measurements.

5. SOME EXTENSIONS

In this section, we present some applications of our technique, to three different flux-splitting solvers of homogeneous shallow water equations.

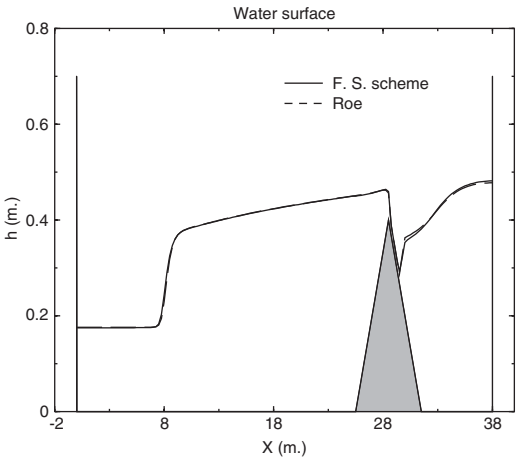

Figure 18. Test 4.2: computed water surface at $t = 20$ s.

We introduce a scheme of first order, another of second order and a third one combining the preceeding two by means of flux limiters. We give a common structure to extend these solvers for non-homogeneous shallow water equations in order to verify the Bermúdez–Vázquez condition.

We propose a new flux-splitting method which is defined by (11), (12) and

$$B_1(W_i, W_{i+1}) = \frac{A(W_i) + |A((W_i + W_{i+1})/2)|}{2},$$
$$B_2(W_i, W_{i+1}) = \frac{A(W_{i+1}) - |A((W_i + W_{i+1})/2)|}{2} \tag{47}$$

This scheme uses a combination different from the schemes of Steger Warming and Vijayasundaram, so that it uses centred evaluations of the upwinding components and decentered evaluations of the centred components. It is a new flux splitting method which exhibits better stability properties than directly Steger Warming or Vijayasundaram methods, in the neighbourhood of sonic points and shock waves.

We construct the extension of this scheme to non-homogeneous shallow water equations. We decompose the source term as

$$G(x_i, W_i) = \tfrac{1}{2}[I + |\tilde{A}^*|\tilde{A}^{-1}]G(x_i, W_i) + \tfrac{1}{2}[I - |\tilde{A}^*|\tilde{A}^{-1}]G(x, W) \tag{48}$$

where $\tilde{A} = \tilde{X}\Lambda\tilde{X}^{-1}$ with $\tilde{X} = PX$ and P is the same eigenvectors scaling matrix as in previous sections.

Notice that in (48) we introduce matrix $\tilde{A}^*$, defined as $\tilde{A}^* = Y\Lambda Y^{-1}$ with $Y = QX$. In previous sections Q is equal to P, but they can be different, in such a way the technique can be applied to other schemes such as the ones we introduce here.

Then, if Q_1 is the scaling matrix related to G_1, Q_3 to G_3 and Q_2^1, Q_2^2 to the first and second component of G_2, we define

$$Q_1 = Q_2^1 = Q_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix} \quad \text{and} \quad P_2^2 = Q_2^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{49}$$

Therefore, the numerical source terms are

$$\mathscr{G}_j^l(x_{i-1}, x_i, x_{i+1}, W_{i-1}, W_i, W_{i+1}) = (\mathscr{G}_j^l)_{\mathrm{L}}(x_{i-1}, x_i, W_{i-1}, W_i) + (\mathscr{G}_j^l)_{\mathrm{R}}(x_i, x_{i+1}, W_i, W_{i+1}) \tag{50}$$

for $j = 1, 2, 3$ and $l = 1, 2$, with

$$(\mathscr{G}_j^l)_{\mathrm{L}}(x_{i-1}, x_i, W_{i-1}, W_i) = \frac{(x_i - x_{i-1})/2}{\Delta x_i}[I + |(B_j^*)^l(W_{i-1}, W_i)|(B_j^l)^{-1}(W_{i-1}, W_i)]\tilde{G}_j(x_{i-1}, x_i, W_{i-1}, W_i) \tag{51}$$

$$(\mathscr{G}_j^l)_{\mathrm{R}}(x_i, x_{i+1}, W_i, W_{i+1}) = \frac{(x_{i+1} - x_i)/2}{\Delta x_i}[I - |(B_j^*)^l(W_i, W_{i+1})|(B_j^l)^{-1}(W_i, W_{i+1})]\tilde{G}_j(x_i, x_{i+1}, W_i, W_{i+1}) \tag{52}$$

where $B_j^l(W_i, W_{i+1})$ is defined as in Section 2.2 and $(B_j^*)^l$ is defined as follows:

$$(B_j^*)^l(W_i, W_{i+1}) = (\tilde{A}^*)_j^l\left(\frac{W_i + W_{i+1}}{2}\right) \quad \text{and} \quad (\tilde{A}^*)_j^l = Y_j^l \Lambda (Y_j^l)^{-1}$$

with $Y_j^l = Q_j^l X$, where we understood $Q_1^1 = Q_1^2 = Q_1$ and $Q_3^1 = Q_3^2 = Q_3$. Then, we have the following result.

Theorem 3
The flux-splitting scheme defined by (11), (12), (47), (49), (50), (51) and (52) verifies the following:

1. The flux-splitting scheme satisfies the exact Bermúdez–Vázquez condition with respect to the stationary solution (24).
2. In the absence of friction terms, the scheme satisfies the exact Bermúdez–Vázquez condition with respect to the stationary solution (43).
3. For general source terms, the scheme satisfies the approximated Bermúdez–Vázquez condition with respect to the stationary solution (43).

Now, we present a new flux-splitting second-order method which is defined by (11), (12) and

$$B_1(W_i, W_{i+1}) = \frac{A(W_i) + (\Delta t/\Delta x_i)A^2((W_i + W_{i+1})/2)}{2}$$
$$B_2(W_i, W_{i+1}) = \frac{A(W_{i+1}) - (\Delta t/\Delta x_i)A^2((W_i + W_{i+1})/2)}{2} \tag{53}$$

This is an extension of Lax–Wendroff's method for hyperbolic equations with constant coefficients. The numerical source term is defined by (50), (51) and (52) where only differ the definition of $(B_j^*)^l$:

$$(B_j^*)^l = \frac{\Delta t}{\Delta x_i}((\tilde{A}^*)_j^l)^2$$

Then, for this scheme we have the same results as in Theorem 3. This is a second-order scheme, therefore this scheme can be oscillatory near shock waves [15]. To avoid this difficulty, we define a new flux-splitting solver which involves the two above schemes by means of flux limiters. This scheme is defined by (11), (12) and

$$B_1(W_i, W_{i+1}) = \frac{A(W_i) + (\psi(r)(\Delta t/\Delta x_i)A^2 + (1 - \psi(r))|A|))((W_i + W_{i+1})/2)}{2}$$
$$B_2(W_i, W_{i+1}) = \frac{A(W_{i+1}) - (\psi(r)(\Delta t/\Delta x_i)A^2 + (1 - \psi(r))|A|))((W_i + W_{i+1})/2)}{2} \tag{54}$$

where $\psi(r)$ is a flux limiter (see Reference [10]). The numerical source term of this scheme is defined by (50)–(52) and

$$(B_j^*)^l = \psi(r)\frac{\Delta t}{\Delta x_i}((\tilde{A}^*)_j^l)^2 + (1 - \psi(r))|(\tilde{A}^*)_j^l|$$

Still in this case Theorem 3 holds.

The technique that we present in this work can also be applied to other schemes, such Steger-Warming or Vijayasundaram. Also, we can apply this technique to other equations, for example, non-homogeneous Euler equations.

6. CONCLUSION

In this paper, we have introduced a flux-splitting solver for 1D shallow water equations with source terms, which satisfies a strengthened consistency condition for stationary solutions. The main methodological innovation is the construction of the numerical source terms using a re-scaling of the eigenvectors of the flux matrix. The practical performances of this solver in several tests cases are quite close to those of Roe's solver, with slightly higher numerical diffusion and quite similar computing cost. We have, thus, developed an easy-to-programme and performing solver for the focused equations. The extension of this solver to a wide class of solvers satisfying the enhanced consistency conditions, of which both Q-schemes and flux-splitting schemes are particular cases, in addition to some new solvers, is in progress and shall be reported in a forthcoming paper.

ACKNOWLEDGEMENTS

This research was partially supported by Spanish Government Research Projects REN2000-1162-C02-01 and REN2009-1168-C02-01. The authors wish to thank Professors Pilar Brufau and Pilar García Navarro for their technical help, and also Professors Antonio Domínguez Delgado, Bosco García Archilla and María Elena Vázquez Cendón for their valuable remarks and in general their interest in the development of this work.

REFERENCES

1. Aves MA, Griffiths DF, Higham DJ. Runge–Kutta solutions of a hyperbolic conservation law with source term. *SIAM Journal on Scientific Computing* 2000; **22**(1):20–38.
2. Burguete J, Garcia-Navarro P. Efficient construction of high-resolution TVD conservative schemes for equations with source terms. Application to shallow water flows. *International Journal for Numerical Methods in Fluids* 2001; **37**:209–248.
3. LeVeque HC, Yee. A study of numerical methods for hyperbolic conservation laws with stiff source terms. *Journal of Computational Physics* 1990; **86**:187–210.
4. Zhou JG, Causon DM, Mingham CG, Ingram D. The surface gradiente method for the treatment of source terms in the shallow-water equations. *Journal of Computational Physics* 2001; **168**:1–25.
5. Bermúdez A, Vázquez Cendón ME. Upwind methods for hyperbolic conservation laws with source terms. *Computers Fluids* 1994; **23**(8):1049–1071.
6. Vázquez Cendon ME. Estudio de esquemas descentrados para su aplicacion a las leyes de conservación hiperbólicas con términos fuente. *Ph.D. Thesis*, Universidad de Santiago de Compostela, 1994.
7. Toro EF. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer: Berlin, 1997.
8. Godlewski E, Raviart PA. Hyperbolic systems of conservation laws. *Mathematiques et Applications*. Ellipses: Paris, 1991.
9. Vijayasundaram G. Resolution Numérique des équations d'Euler pour des écoulements transsoniques avec un schéma de Godunov en éléments finis. *Ph.D. thesis*, L'Université Pierre et Marie Curie Paris, Vol. VI, 1982.
10. Godlewski E, Raviart PA. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*. Springer: Berlin, 1996.
11. Bermudez A, Dervieux A, Desideri JA, Vázquez Cendón ME. Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes. *Computer Methods in Applied Mechanics and Engineering* 1998; **155**:49.
12. Brufau P. Simulación bidimensional de flujos hidrodinámicos transitorios en gemotrías irregulares. *Ph.D. Thesis*, Universidad de Zaragoza, 2000.
13. Maria Elena Vázquez Cendón. Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry. *Journal of Computational Physics* 1999; **148**:497–526.
14. Chacón T, Franco D, Ortegón F, Sánchez I. Modelling of compressible flows with highly oscillating initial data by homogenization. *Applied Numerical Mathematics* 1998; **26**:435–464.
15. Alcrudo F, García-Navarro P. A high-resolution Godunov-type scheme in finite volumes for the 2D shallow-water equations. *International Journal for Numerical Methods in Fluids* 1993; **16**:489–505.

