XXX° Convegno di Idraulica e Costruzioni Idrauliche - IDRA 2006
MASTER CLASS *Modelli numerici di correnti fluviali su fondo fisso e fondo mobile*

A MARCHING IN SPACE AND TIME (MAST) SOLVER OF THE SHALLOW WATER EQUATIONS

Costanza Aricò

Dipartimento di Ingegneria Idraulica ed Applicazioni Ambientali, Università di Palermo – Palermo (IT)
email arico@idra.unipa.it

Keywords: shallow water, unsteady flow, numerical methods, Eulerian methods, fractional time-step methods, flow routing

ABSTRACT

A new approach is presented for the numerical solution of the complete 1D and 2D Saint-Venant equations. At each time step, the governing system of Partial Differential Equations (PDEs) is split, using a fractional time step methodology, into a convective prediction system and a diffusive correction system. Convective prediction system is further split into a convective prediction and a convective correction system, according to a specified approximated potential. If a scalar exact potential of the flow field exists, correction vanishes and the solution of the convective correction system is the same solution of the prediction system. A MArching in Space and Time (MAST) technique is used for the solution of the two systems. MAST solves a system of two Ordinary Differential Equations (ODEs) in each computational cell, using for the time discretization a self-adjusting fraction of the original time step. The computational cells are ordered and solved according to the decreasing value of the potential in the convective prediction step and to the increasing value of the same potential in the convective correction step. The diffusive correction system is solved using an implicit scheme, that leads to the solution of a large linear system, with the same order of the cell number, but sparse, symmetric and well conditioned. The numerical model shows unconditional stability with regard of the Courant number (CFL), requires no special treatment of the source terms and a computational effort almost proportional to the cell number. Several tests have been carried out using analytical solutions, as well as experimental data.

1 INTRODUCTION

Saint-Venant (SV) equations (or shallow water equations) are commonly applied for the simulation of unsteady shallow water flows (*De Saint Venant*, 1871). Both explicit and implicit numerical methods have been extensively used for the computation of their numerical solution. Even though implicit schemes are stable for CFL numbers greater than 1, they lead to the solution of large non linear algebraic systems, and the computational effort grows very quickly along with the number of elements. Implicit Preissman schemes show also other limitations, specially for simulations of transcritical flows (*Meselhe et al., 1997*). For these reasons, most of the algorithms recently proposed by the Authors fall in the class of explicit methods.

Several Godunov-type schemes have been proposed for the "non viscous" form of the SV equations, where source term in the momentum equation is zero (*Alcrudo and Garcìa-Navarro*, 1993; *Mingham and Causon*, 1998). A fractional time step procedure (*Toro*, 1997) can be applied for the solution of the shallow water equations using Godunov-type schemes when source terms cannot be neglected. The corresponding non-homogeneous form of the governing equations is solved in two sequential steps. During the first step, the homogeneous problem (without the source term) is solved. In the second step, a set of ODEs systems including source terms are solved sequentially, one after the other. A similar approach is simple, but often produces poor solutions, especially in stationary or quasi-stationary cases (*LeVeque*, 1998). In fact, when the local time derivative of the flow variables is negligible, the spatial

derivative of the flux terms should exactly balance the source term. Fractional step approach in this case can fail, since the solution of the homogeneous problem may lead to large changes in the independent variables, which are difficult to correct by solving the ordinary differential equations during the second step.

In the last decade many Authors have developed Godunov-type schemes for the solution of balance laws, including source terms. The main focus of these schemes is the balance between the numerical flux and the source term, when the same numerical discretization is applied. *Bermudez and Vazquez-Cendon* (1994), using a three-points upwind scheme with 1st spatial approximation order and upwind discretization of both flux and source terms, introduced the definition of "conservation property". This property is the accuracy provided by a numerical scheme in the solution of the steady-state problem representing water at rest (constant total head and zero flow rate in all the domain). If the conservative property is not satisfied, exactly or approximately, the propagation of unphysical oscillation is detected also in non-stationary problems. In the paper of *Bermudez and Vazquez-Cendon* (1994) source term is given only by the bottom slope in hydrostatic condition. *Vazquez-Cendon* (1999) extended the definition of conservation property to the case of uniform flow in rectangular sections. *Hubbard and Garcìa-Navarro* (2000) proved this conservation property for the 2nd order MUSCL scheme with flux limiter.

Upwinding the source term is computationally expensive, because source term has to be projected on an eigenvectors basis. *LeVeque* (1998) developed an approach where the source term is embedded into the wave-propagation algorithm, avoiding the fractional step inconvenient mentioned before. The Author defined a Riemann problem inside the cell to balance the source term and the flux gradient; the method preserves steady and quasi-steady flow conditions.

Starting from the work of *Bermudez and Vazquez-Cendon* (1994), *Vukovic and Sopta* (2002) and *Zic et al.* (2004) developed an early version of the higher order schemes (ENO and WENO), where numerical flux terms exactly balance the source terms. The main limitation of these numerical methods proposed for balance laws is that they have been developed for specific form of the source terms and miss generality.

Zhou et al. (2001) introduced the surface gradient method (SGM) for the accurate treatment of the bed slope terms in the shallow waters equations. The Authors proposed a Godunov-type scheme where, in contrast to conventional data reconstruction methods based on the conservative variables, the water surface level is chosen as the basis for data reconstruction. The conservative variables are accurately reconstructed by a 2nd order scheme at cell interfaces and the numerical fluxes are computed by a Riemann solver. Integration in time is performed by means of a predictor and a corrector step. More recently, the SGM has been used by the same Authors (*Zhou et al.*, 2002) to deal with bed topography with vertical steps. The method is referred to as the surface gradient method for steps (SGMS). Both SGM and SGMS produce accurate solution over structured meshes.

All these methods have been proposed for structured meshes and are restricted in their ability to fit irregular and/or curvilinear boundaries. A popular approach for the treatment of irregular boundaries is the use of boundary-fitted grids which make the boundary contour a coordinate surface (*Hauser et al*., 1985; *Yang and Hsu*, 1993). The flow equations are transformed from Cartesian to curvilinear coordinates and then approximated using finite differences. A grid can be generated iteratively by means of an elliptic solver (*Thompson et al*., 1974) or directly by an algebraic technique such as transfinite interpolation (*Gordon et al*., 1973). The complexity of flow and momentum equations written in the transformed coordinates increases and discretization of the metrics transformation derivatives could produce new approximation errors.

Alternatively to the boundary-fitted grids, *Causon et al*. (2000) extended the Cartesian cut cell approach (*Berger et al*., 1995; *De Zeeuw and Powell*, 1993) to a Godunov-type scheme based on a MUSCL reconstruction and suitable approximate Riemann solver. In the Cartesian cut cell approach, boundary contours are cut out of a background Cartesian mesh and cells partially or completely cut are signed out for special treatment in the computation of the gradients and reconstruction of the flow data (*Causon et al*., 2000). Since a cut cell may turn out to be arbitrarily small, numerical stability requirements may force a very small time step. To overcome this problem, a cell merging technique is implemented (*Chiang et al*., 1992). The basic idea is to combine several neighbouring cells together so that interfaces between merged cells are ignored and waves travel in a new combined larger cell without reducing the global time step. The remainder of the flow cells are treated in a straightforward manner.

Successively, *Zhou et al*. (2004) applied the Cartesian cut cell approach to the SGM and SGMS methods.

Triangular mesh is generally the simplest and most convenient method for covering a 2D domain. An advantage of using triangular meshes is their ability to generate grids on arbitrary geometries and to increase the number of cells in high-gradient regions or in regions of particular interest in the flow field. Recently, several numerical schemes have been proposed to solve the shallow water equations over unstructured meshes. *Anastasiou and Chan* (1997) developed a 2D depth integrated 2nd order Godunov-type scheme, based on a cell-centred finite volume upwind formulation. The Roe's flux function is used for the evaluation of the inviscid fluxes at cell interfaces, solving a Riemann problem in the direction normal to the cell interface. The viscous terms are computed using a 2nd order accurate finite volume formulation. Time integration is done according to a 2nd order trapezoidal implicit scheme.

Hubbard (1999) proposed a 2nd order MUSCL scheme over unstructured meshes. The Author solves the homogeneous form of the shallow water equations and applies the Roe's approximate Riemann solver at cell interfaces for the flux estimation. The Author uses a 2nd order Runge-Kutta time stepping for time integration.

Jiwen and Ruxun (2001) applied a composite finite volume formulation for the solution of the homogeneous SV equations. The Authors combine a Lax-Wendroff and a Lax-Friedrich scheme into a multi-step composite scheme. The Lax-Wendroff scheme is highly dispersive and produces unphysical oscillations; on the opposite, the Lax-Friedrich scheme is characterized by numerical diffusion. The composite scheme combines these two methods in a step-by-step scheme to exploit their merits and remove their deficiencies.

Yoon and Kang (2004) propose a 2nd order finite volume scheme over unstructured triangular mesh. The Authors apply the HLL approximate Riemann solver for flux estimation at the cell boundaries. To circumvent numerical instabilities, the Authors split the source term in the friction and the bottom slope components applying an operator splitting technique. Time integration is made by a 3rd order TVD-Runge-Kutta method (*Shu and Osher*, 1988).

In these Godunov-type schemes time step size is limited by the Courant-Friedrichs-Levy (CFL) condition.

In the present paper we further developed the recent MAST algorithm for the solution of the 1D and 2D SV equations. MAST has been already developed for the solution of the simple convective transport problem (*Bascià and Tucciarelli*, 2004; *Aricò and Tucciarelli*, 2006), the 1D and 2D shallow water problem with irrotational flow field (*Tucciarelli and Termini*, 2000; *Noto and Tucciarelli*, 2007), as well as in the 1D shallow water problem with a mono-oriented flow (*Tucciarelli*, 2003). The methodology is not restricted by the CFL condition in the choice of the time step and, even if the solution of a large linear system is required for each time step, the corresponding computational cost is negligible and the total computational cost is almost proportional to the cell number.

The governing equation system is split first in a convective component and in a diffusive component. The convective system is solved applying a MArching in Space and Time (MAST) procedure, where the convective fluxes are solved using an Eulerian approach and the computational cells are required to be ordered and solved according to a decreasing scalar potential value.

The correction step computes the diffusive corrective fluxes by means of the solution of a large linear system, that has the order of the cell number, but is sparse, symmetric and well conditioned. The potential is defined as a scalar whose gradient is opposite to the flux direction. Because the flow field computed solving the SV equations has not a scalar potential, an approximated potential has to be used, along with a second correction step that is aimed to compute the rotational fluxes. This approximated potential is defined according to the flow direction inside each element at the beginning of each time step. In the following of the paper, the application of the numerical procedure to the 1D case will be shown first and then to the 2D case.

2 THE MAST PROCEDURE IN THE CASE OF 1D MONO-ORIENTED FLOW

The 1D SV equations in a channel with non-prismatic section can be written in the following form (figure 1) (see for example *Abbott and Minns*, 1992):

$$\frac{\partial \sigma}{\partial t} + \frac{\partial q}{\partial x} = 0 \tag{1},$$

$$\frac{\partial q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q^2}{\sigma}\right) + g\sigma\frac{\partial h}{\partial x} + g\sigma\left(\frac{n^2 q|q|}{\sigma^2 R^{4/3}} - S_0\right) = 0 \tag{2},$$

where x is the flow direction, t is the time, g is the gravitational acceleration, h is the water depth, σ is the flow cross section, q is the stream flow rate, S_0 is the bottom bed slope ($S_0 = -\frac{\partial z}{\partial x}$ where z is the bottom elevation with respect to a reference level), R is the hydraulic radius and n is the Manning friction coefficient. Eq. (2) can also be written in the following form:

$$\frac{\partial q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q^2}{\sigma}\right) + g\sigma\frac{\partial H}{\partial x} + g\frac{n^2 q|q|}{\sigma R^{4/3}} = 0 \tag{3},$$

where H is the total water level, $H = h + z$. The unknown variables in Eqq. (1) and (3) are the water cross section $\sigma(h)$ and the flow rate q.

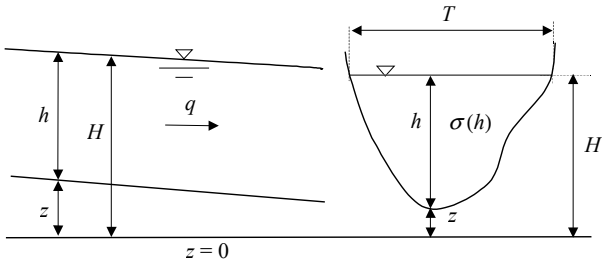

Figure 1. Definition sketch

Assume the following general system of balance equations:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial \mathbf{F}(\mathbf{U})}{\partial x} = \mathbf{B}(\mathbf{U}) \tag{4},$$

where **U** is the unknowns vector, **F**(**U**) is the numerical flux-vector term and **B**(**U**) is the source term. Using a fractional time step strategy, we set:

$$\mathbf{F}(\mathbf{U}) = \mathbf{F}^p(\mathbf{U}) + \left(\mathbf{F}(\mathbf{U}) - \mathbf{F}^p(\mathbf{U})\right), \quad \mathbf{B}(\mathbf{U}) = \mathbf{B}^p(\mathbf{U}) + \left(\mathbf{B}(\mathbf{U}) - \mathbf{B}^p(\mathbf{U})\right) \tag{5},$$

where $\mathbf{F}^p(\mathbf{U})$ and $\mathbf{B}^p(\mathbf{U})$ are suitable numerical flux-vector and source terms, further defined.

After integration in time, system (4) can be split in the two following systems:

$$\mathbf{U}^{k+1/2} - \mathbf{U}^k + \frac{\partial}{\partial x}\int_0^{\Delta t} \mathbf{F}^p dt = \int_0^{\Delta t} \mathbf{B}^p dt \tag{6,a},$$

$$\mathbf{U}^{k+1} - \mathbf{U}^{k+1/2} + \frac{\partial}{\partial x}\int_0^{\Delta t} \mathbf{F}\, dt - \overline{\mathbf{F}}^p = \int_0^{\Delta t} \mathbf{B}\, dt - \overline{\mathbf{B}}^p \tag{6,b},$$

where $\overline{\mathbf{F}}^p$ and $\overline{\mathbf{B}}^p$ are the mean values of the flux and the source terms as computed in the prediction

step. We call systems (6,a)-(6,b) respectively prediction and correction system. $\mathbf{U}^{k+1/2}$ and $\mathbf{U}^{k+1}$ are the unknown variables computed at the end of respectively the prediction and the correction step. Observe that, summing Eqq. (6,a) and (6,b) the integral of the original system (4) is formally obtained. The numerical corrected solution will be close to the solution of this one as far as the difference between the flux and the source terms is either small or time-independent. The advantage of using formulation (6) instead of (4) is that, with a suitable choice of the prediction terms $\mathbf{F}^p(\mathbf{U})$ and $\mathbf{B}^p(\mathbf{U})$, each of the two systems (6,a)-(6,b) can be much easier to solve than the original system (4).

For the SV equations solution, we set:

$$\mathbf{U}_1 = \sigma, \quad \mathbf{U}_2 = q \tag{7,a},$$

$$\mathrm{F}_1 = q\,, \quad \mathrm{F}_2 = \frac{q^2}{\sigma} + gM\,, \quad \mathrm{B}_1 = 0\,, \quad \mathrm{B}_2 = g\sigma S_0 - \frac{gn^2q^2}{\sigma R^{4/3}} \tag{7,b},$$

$$\mathrm{F}_1^p = q\,, \quad \mathrm{F}_2^p = \frac{q^2}{\sigma}\,, \quad \mathrm{B}_1^p = 0\,, \quad \mathrm{B}_2^p = -g\sigma\frac{\partial H^k}{\partial x} - \frac{gn^2q^2}{\sigma R^{4/3}} \tag{7,c},$$

where M is the momentum of the cross section with respect to its top width. The total head gradient in the B_2^p term is computed at time level k and is kept constant along the time step. For sake of simplicity, assume a mono-oriented positive velocity and a cell numeration increasing from 1 to N in the downstream direction (1, 2,…, N). The argument of the integral prediction system (6,a) is given by:

$$\begin{aligned} &\frac{\partial\sigma}{\partial t} + \frac{\partial q}{\partial x} = 0 \\ &\frac{\partial q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q^2}{\sigma}\right) = \mathrm{B}_2^p \end{aligned} \tag{8},$$

that differs from the original one only in the time level of the total head gradient included in the source term B_2^p. The argument of the integral correction system (6,b) is:

$$\begin{aligned} &\frac{\partial\sigma}{\partial t} + \frac{\partial q}{\partial x} = \frac{\partial\overline{\mathrm{F}}_1^p}{\partial x} \\ &\frac{\partial q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{q^2}{\sigma} + gM\right) - \frac{\partial\overline{\mathrm{F}}_2^p}{\partial x} = g\sigma S_0 - \frac{gn^2q^2}{\sigma R^{4/3}} - \overline{\mathrm{B}}_2^p \end{aligned} \tag{9}.$$

Observe that the difference between the flux gradient and the source term goes to zero, in system (9), along with the time distance with respect to the initial time level t^k. This implies that both the unknown changes and the relative errors go to zero along with the time step size.

In system (9) we assume the difference between the gradient of the momentum flux $\frac{\partial}{\partial x}\left(\frac{q^2}{\sigma}\right)$ and the gradient of the mean value of the corresponding term computed in the prediction step $\frac{\partial\overline{\mathrm{F}}_2^p}{\partial x}$ to be negligible.

According to this hypothesis, after simple manipulations, systems (8) and (9) can be written in quasi-linear matrix form (Tucciarelli, 2003):

$$\frac{\partial \mathbf{U}}{\partial t}+\mathbf{A}_1\begin{pmatrix}\dfrac{\partial \sigma}{\partial x}\\ \dfrac{\partial q}{\partial x}\end{pmatrix}=\begin{pmatrix}0\\ \mathrm{B}_2^p\end{pmatrix}$$
$$\frac{\partial \mathbf{U}}{\partial t}+\mathbf{A}_2\begin{pmatrix}\dfrac{\partial \sigma}{\partial x}\\ \dfrac{\partial q}{\partial x}\end{pmatrix}=\begin{pmatrix}\dfrac{\partial \overline{\mathrm{F}}_1^p}{\partial x}\\ g\sigma S_0-\dfrac{gn^2q^2}{\sigma R^{4/3}}-\overline{\mathrm{B}}_2^p\end{pmatrix} \quad (10),$$

where the Jacobians of systems (10) are respectively:

$$\mathbf{A}_1=\begin{pmatrix}0 & 1\\ -\dfrac{q^2}{\sigma^2} & \dfrac{2q}{\sigma}\end{pmatrix}, \qquad \mathbf{A}_2=\begin{pmatrix}0 & 1\\ g\overline{h} & 0\end{pmatrix} \quad (11),$$

$\overline{h}=\sigma/T$ and T is the cross section top width. It is easy to show that $\mathbf{A}_1$ eigenvalues are $\lambda_1 = \lambda_2 = ve$, where ve is the vertically averaged flow velocity. Observe that the solution of system (8) is equivalent to the solution of a single non-linear convection equation, function of the head gradient at time level t^k. $\mathbf{A}_2$ eigenvalues are $\lambda_1 = \sqrt{g\overline{h}}$, $\lambda_2 = -\lambda_1 = -\sqrt{g\overline{h}}$. This implies that in the correction system convective fluxes are missing. For these reasons we call the prediction and the correction systems respectively convective prediction system and diffusive correction system.

Define $\mathbf{B}^p(\mathbf{U}_j)$ the prediction source terms computed as function of the unknowns $\mathbf{U}_j = (\sigma_j, q_j)$ in cell j and $\mathbf{F}^p(\mathbf{U}_j, \mathbf{U}_{j+1})$ the prediction flux between cell j and j+1.

We adopt a first order spatial approximation of the unknown variables inside the cells. After integration in space and time, Eqq. (10) can be written as:

1) convective prediction system:

$$\frac{\sigma_j^{k+1/2}-\sigma_j^k}{\Delta t}\Delta x+\frac{1}{\Delta t}\int_0^{\Delta t}\left(q_j-q_{j-1}\right)dt=0, \quad j=1,\ldots,N \quad (12,a),$$

$$\frac{q_j^{k+1/2}-q_j^k}{\Delta t}\Delta x+\frac{1}{\Delta t}\int_0^{\Delta t} f_{j,j+1}\frac{q_j}{\sigma_j}dt-\frac{1}{\Delta t}\int_0^{\Delta t} f_{j-1,j}\frac{q_{j-1}}{\sigma_{j-1}}dt=\frac{1}{\Delta t}\int_0^{\Delta t}\mathrm{B}_2^p\left(\mathbf{U}_j\right)\Delta x\,dt, \quad j=1,\ldots,N \quad (12,b);$$

2) diffusive correction system:

$$\frac{\sigma_j^{k+1}-\sigma_j^{k+1/2}}{\Delta t}\Delta x+\frac{1}{\Delta t}\int_0^{\Delta t}\left(q_j-q_{j-1}\right)dt=\overline{F}_1^p\left(\mathbf{U}_j,\mathbf{U}_{j+1}\right)-\overline{F}_1^p\left(\mathbf{U}_{j-1},\mathbf{U}_j\right) \quad (13,a),$$

$$\frac{q_j^{k+1}-q_j^{k+1/2}}{\Delta t}\Delta x+\frac{g}{\Delta t}\int_0^{\Delta t}\left(M_{j,j+1}-M_{j-1,j}\right)dt=\frac{1}{\Delta t}\int_0^{\Delta t}\left(g\sigma_j S_0-\frac{gn^2q_j^2}{\sigma_j R_j^{4/3}}\right)\Delta x dt-\overline{\mathrm{B}}_2^p\Delta x \quad (13,b),$$

where $M_{j,j+1}$ is the cross section momentum at the interface between cells j and j+1.

2.1 Solution of the convective prediction system

Convective prediction system is solved using a MArching in Space and Time (MAST) procedure. Because of the sign of the eigenvalues of the Jacobian $\mathbf{A}_1$, it is possible to solve each Eqq. (12) if a polynomial approximation of the volume and the momentum fluxes entering from the upstream cell is always known. This condition can be guaranteed if the flow direction has a constant orientation or, more generally, if a scalar potential exists. In this case, it is always possible to order and solve sequentially the cells one after the other. In the case of mono-oriented flow, assuming a cell index increasing in the downstream direction, Eqq. (12) can be solved as a standard ODEs system, integrated from time 0 to time Δt, that is:

$$\frac{d\sigma}{dt}\Delta x + q_j - \phi_{j-1,j}(t) = 0 \qquad (14,a),$$

$$\frac{dq}{dt}\Delta x + f_{j,j+1}\frac{q_j}{\sigma_j} + \left(\frac{n^2 q_j |q_j| g}{\sigma_j R_j^{4/3}} + g\sigma_j \frac{\partial H_j^k}{\partial x}\right)\Delta x - \xi_{j-1,j}(t) = 0 \qquad (14,b),$$

In Eqq. (14), the terms $\phi_{j-1,j}(t)$ and $\xi_{j-1,j}(t)$ are a polynomial (known) approximation of respectively the volume and the momentum flux entering cell j from cell j-1. To solve Eqq. (14) a Runge-Kutta method, with self-adapting time step, is applied (*Nag Library Manual*, 2005). After the ODEs in cell j are solved, polynomial approximations required in cell j+1 can be computed as approximation of the volume and the momentum fluxes leaving cell j along time step Δt. The basic idea of the numerical technique is to compute the solution, within a given time step, by marching in space along the flux direction through all the computational domain. The mass and momentum fluxes entering the first cell are given by the upstream boundary conditions, as better specified in a following section.

Observe that it is possible to solve sequentially the ODEs systems, without any restriction on the size of the time step, because the characteristic line of the problem (8) is oriented according to the velocity direction and the velocity sign has been assumed to be constant. This is equivalent, after solution of cell 1 from time t^k to time t^{k+1}, to "translate" the boundary condition (thick segment in figure 2) at the interface between cell 1 and 2, to get the initial value of the characteristic curve carrying on the solution of cell 2 and so on through all the computational domain. Because of this, the method can be classified, other than "marching in time", also "marching in space".

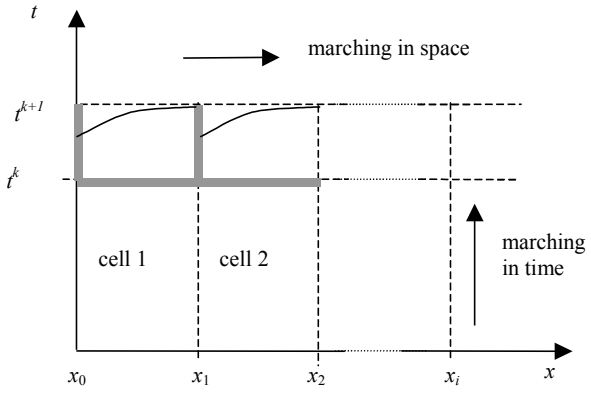

Figure 2. The core of the marching in space and time procedure

2.1.1 Polynomial time approximation

A 3rd order approximating in time polynomial has been chosen for the entering volume and momentum fluxes. The corresponding coefficients are obtained by setting the initial, the final and the

average values of the polynomials equals to the computed ones. This implies:

$$\phi^0_{j,j+1} = q^k_j, \quad \phi^0_{j,j+1} + \phi^1_{j,j+1}\Delta t + \phi^2_{j,j+1}\Delta t^2 = q^{k+1/2}_j \tag{15,a),}$$

$$\phi^0_{j,j+1} + \phi^1_{j,j+1}\frac{\Delta t}{2} + \phi^2_{j,j+1}\frac{\Delta t^2}{3} = \overline{f}_{j,j+1} \tag{15,b),}$$

$$\xi^0_{j,j+1} = m^k_{j,j+1}, \quad \xi^0_{j,j+1} + \xi^1_{j,j+1}\Delta t + \xi^2_{j,j+1}\Delta t^2 = m^{k+1/2}_{j,j+1} \tag{16,a),}$$

$$\xi^0_{j,j+1} + \xi^1_{j,j+1}\frac{\Delta t}{2} + \xi^2_{j,j+1}\frac{\Delta t^2}{3} = \overline{m}_{j,j+1} \tag{16,b),}$$

where

$$m_{j,j+1} = q^2_j / \sigma_j \tag{17).}$$

The average leaving volume flux $\overline{f}_{j,j+1}$ is computed by applying the cell mass balance, that is:

$$\overline{f}_{j,j+1} = \overline{f}_{j-1,j} - \frac{\Delta x\left(\sigma^{k+1/2}_j - \sigma^k_j\right)}{\Delta t} \tag{18).}$$

Due to the non-linearity of the momentum balance equation, the average value of the leaving momentum flux $\overline{m}_{j,j+1}$ is computed via numerical time integration of the values computed by Eq. (17). The solution at the Gauss points is obtained by a C^1 interpolation of the solution values produced by the Runge-Kutta method used for the solution of the ODEs system.

2.1.2 Total head gradient computation

In the convective prediction system the total head gradient of cell *j* at the interface with cell *j* + 1 downstream along the flow direction at the beginning of each time step is computed as:

$$\frac{\partial H^k_j}{\partial x} \cong \frac{H^k_{j+1} - H^k_j}{\Delta x} \qquad j = 1, \ldots, N\text{-}1 \tag{19).}$$

In the last element, according to the hypothesis of constant spatial head gradient along the time step, it is assumed the head change to be equal to the upstream element one. From this assumption the following relationships hold:

$$\sigma\left(h^{k+1/2}_N\right) = \sigma\left(h^{k+1/2}_{N-1}\right), \quad f^{k+1/2}_N = q^{k+1/2}_{N-1} - \frac{\sigma^{k+1/2}_N - \sigma^k_N}{\Delta t}\Delta x \tag{20),}$$

where $f^{k+1/2}_N$ is the downstream boundary flux predicted at the end of the time step.

2.1.3 Small water depths computation

The friction and the momentum flux terms estimation, in the momentum balance equation, is subject to large errors when water depth is close to zero. This uncertainty affects the result stability mainly in the decreasing part of the hydrograph, because the small water depth values tend to remain and further reduce for a large computational time. To deal with this well-known problem, the following procedure has been applied.

A minimum water depth value h_m, negligible for practical applications, is first selected. Call σ_m the corresponding cross section. We assume the Manning relationship and the momentum flux definition to hold only for $h > h_m$. In the other case, these two terms are replaced with the following fictitious ones (see also figure 3):

$$\frac{n^2 q|q|}{\sigma R^{4/3}} = n^2 q|q| P^{4/3}\left(\frac{1}{\sigma_m^{5/3}} + (\sigma-\sigma_m)\frac{\partial f(\sigma)}{\partial \sigma}\bigg|_{\sigma=\sigma_m}\right) \tag{21},$$

$$\frac{q^2}{\sigma} = q^2\left(\frac{1}{\sigma_m} + (\sigma-\sigma_m)\frac{\partial f(\sigma)}{\partial \sigma}\bigg|_{\sigma=\sigma_m}\right) \tag{22},$$

where P in Eq. (21) is the wetted perimeter and function $f(\sigma)$ in Eqq. (21) and (22) is given respectively by:

$$f(\sigma) = \frac{1}{\sigma^{5/3}} \tag{23},$$

$$f(\sigma) = \frac{1}{\sigma} \tag{24}.$$

When the local water depth drops below the minimum value, because of the reduction of the new resistance terms with respect to the real ones, the water depth drops quickly to the zero value at the t^* time. At this point the ODEs system integration is left and the small water volume entering the cell during the remaining $\Delta t - t^*$ time is stored in the same cell to maintain the mass balance.

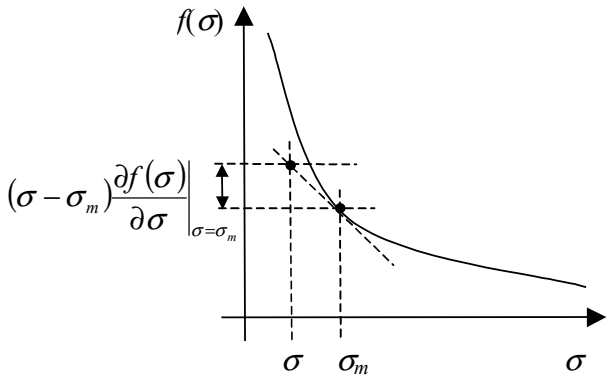

Figure 3. Function approximation for small water depths

2.2 Solution of the diffusive correction system

Diffusive correction system (13) is further simplified and linearized, using a fully implicit finite difference discretization. The resulting system is:

$$T_j^{k+1/2}\frac{H_j^{k+1} - H_j^{k+1/2}}{\Delta t}\Delta x + \left[q_j^{k+1} - q_{j-1}^{k+1} - \left(q_j^{k+1/2} - q_{j-1}^{k+1/2}\right)\right] = 0\,, \qquad j=1,\ldots,N \tag{25,a},$$

$$\frac{q_j^{k+1}-q_j^{k+1/2}}{\Delta t}+gn^2 q_j^{k+1}\left(\frac{|q_j|}{\sigma_j R_j^{4/3}}\right)^{k+\frac{1}{2}}+g\sigma_j^{k+\frac{1}{2}}\frac{\partial H_j^{k+1}}{\partial x}-\left(\frac{n^2 q_j|q_j|g}{\sigma_j R_j^{4/3}}\right)^{k+\frac{1}{2}}-g\sigma_j^{k+\frac{1}{2}}\frac{\partial H_j^{k}}{\partial x}=0 \tag{25,b},$$

where the cross section time derivative discretization has been replaced in Eq. (13,a) with a function of the total head time derivative discretization. Eq. (25,b) can be written in the following form:

$$q_j^{k+1}=-disp_j\frac{\partial H_j^{k+1}}{\partial x}+q_j^{k+1/2}+disp_j\frac{\partial H_j^{k}}{\partial x} \tag{26},$$

where $disp_j=\dfrac{g\sigma_j^{k+1/2}}{\left(\dfrac{1}{\Delta t}+gn^2\left(\dfrac{|q_j|}{\sigma_j R_j^{4/3}}\right)^{k+1/2}\right)}$.

Merging Eq. (26) in Eq. (25,a) we get:

$$T_j^{k+1/2}\frac{H_j^{k+1}-H_j^{k+1/2}}{\Delta t}\Delta x+\left(ds_{j,j+1}^{k+1}-ds_{j-1,j}^{k+1}\right)=\left(ds_{j,j+1}^{k+1/2}-ds_{j-1,j}^{k+1/2}\right),\quad j=1,\ldots,N \tag{27,a},$$

$$ds_{j,j+1}^{k+1}=-disp_j\frac{\partial H_j^{k+1}}{\partial x},\quad ds_{j-1,j}^{k+1}=-disp_{j-1}\frac{\partial H_{j-1}^{k+1}}{\partial x} \tag{27,b}.$$

Using the total head gradient discretization given in Eq. (19), Eqq. (27) are changed into a system of linear equations, with a symmetric, positive defined and well-conditioned matrix. After system (27) is solved in the H^{k+1} unknowns, the flow rates q^{k+1} can be computed by means of Eqq. (26).

2.3 Boundary conditions

Boundary conditions assumed in convective prediction and diffusive correction steps are reported in table 1. h_n e h_c are respectively the boundary (known) water depth and the critical water depth corresponding to the flow rate q_b in the last cell of the domain. The Froude number of the downstream, leaving flux is computed using the domain flow rate and water depth values, the Froude number of the upstream flux is computed using the external flow rate and water depth values. In the convective prediction system both the incoming volume and momentum fluxes are assigned at the upstream cell. If the incoming flux is sub-critical, the momentum flux is computed as function of the water depth in the upstream cell computed at the end of the previous time step. In the case of supercritical downstream flow, solution of the diffusive problem (27) is obtained by setting zero diffusive flux to the downstream boundary (case n. 1, 5 and 6 of table 1). This is equivalent to set:

$$q_b^{k+1}=q_b^{k+1/2} \tag{28,a}.$$

In the cases 2 and 4 of table 1, a water depth corresponding to the known value is assigned:

$$h_b^{k+1}=h_n^{k+1} \tag{28,b}.$$

In the third case of table 1, a critical flow rate corresponding to the computed water depth is finally assigned:

$$q_b^{k+1} = \sigma_b \sqrt{g h_b^{k+1/2}} \qquad (28,c).$$

case n.	Flux	Fr. number	h_n/h_c	Boundary conditions
1	leaving	≥ 1	≤ 1	No bound. cond.
2	leaving	≥ 1	≥ 1	h_n
3	leaving	< 1	< 1	Critical water depth
4	leaving	< 1	≥ 1	h_n
5	incoming	≥ 1		q_n e h_n
6	incoming	< 1		q_n

Table 1. Boundary conditions

3 EXTENSION OF THE MAST PROCEDURE TO THE COMPLETE 1D DYNAMIC SHALLOW WATER EQUATIONS

In the most general case of 1D or 2D flow, the velocity direction and orientation is not known. If the flow field has a scalar potential, like in the case of diffusive models (*Noto and Tucciarelli*, 2001), the flow direction is, at time t^k, opposite to the gradient of the potential computed at the same time and it is possible to order and solve the elements according to the their potential decreasing value. In the most general case, it is still possible to extend the MAST procedure adding a further convective correction step and using an auxiliary scalar function, called approximated potential, further defined.

Call $f_{j,b}$ the flux between cell j and one of the two connected cells ($b = j$-1 or $b = j$+1), taken positive or negative if the flux is respectively leaving or entering cell j and $ve_{j,b}$ the velocity of the corresponding particles, that is:

$$f_{j,b} = q_j(b-j),\ ve_{j,b} = \frac{q_j}{h_j},\ \text{if } q_j(b-j) > 0 \ \text{ and } \ q_j(b-j) > q_b(j-b) \qquad (29,a),$$

$$f_{j,b} = -q_b(j-b),\ \ ve_{j,b} = \frac{q_b}{h_b} \quad \text{otherwise} \qquad (29,b).$$

Observe that, according to the previous rule, condition $f_{j,b} = -f_{b,j}$ always holds.

Assume that a scalar value φ_j^k, called approximated potential, is assigned to each computational cell j at the beginning of each time level t^k. The chosen volume and momentum flux computation rules can guarantee the existence of a potential function that is exact at the beginning of the time step. This approximated potential can be found by setting arbitrarily its value φ_1^k in the first cell and by assigning to the other cells the following values:

$$\varphi_{j+1}^k = \varphi_j^k + 1 \quad \text{if} \quad f_{j,j+1}^k < 0,\ \ \varphi_{j+1}^k = \varphi_j^k - 1 \quad \text{if} \quad f_{j,j+1}^k \geq 0 \qquad \text{for} \quad j = 2,\ldots,N \qquad (30).$$

Even if the flow rates are consistent with the potential gradient at the beginning of the time step, they can change sign along the solution of the convective prediction system (8). For this reason, the solution of a new convective correction system is now required.

System (1)-(3) can be solved by means of the sequential solution of the following three PDEs systems:

1) convective prediction system:

$$\begin{aligned}&\frac{\partial \sigma}{\partial t}+max\left(0,sign\left(-grad\varphi^k q\right)\right)\frac{\partial q}{\partial x}=0\\&\frac{\partial q}{\partial t}+max\left(0,sign\left(-grad\varphi^k q\right)\right)\frac{\partial}{\partial x}\left(\frac{q^2}{\sigma}\right)+g\sigma\frac{\partial H^k}{\partial x}+\frac{gn^2q|q|}{\sigma R^{4/3}}=0\end{aligned} \tag{31},$$

2) convective correction system:

$$\begin{aligned}&\frac{\partial \sigma}{\partial t}-min\left(0,sign\left(-grad\varphi^k q\right)\right)\frac{\partial q}{\partial x}=0\\&\frac{\partial q}{\partial t}-min\left(0,sign\left(-grad\varphi^k q\right)\right)\frac{\partial}{\partial x}\left(\frac{q^2}{\sigma}\right)=0\end{aligned} \tag{32},$$

3) diffusive correction system:

$$\begin{aligned}&\frac{\partial \sigma}{\partial t}+\frac{\partial q}{\partial x}=\frac{\partial q^{k+2/3}}{\partial x}\\&\frac{\partial q}{\partial t}+g\frac{\partial M}{\partial x}+g\sigma^{k+2/3}\left(-S_0+\frac{n^2q|q|}{\sigma^2R^{4/3}}\right)=g\sigma^{k+2/3}\left(\frac{\partial H^k}{\partial x}+\left(\frac{n^2q|q|}{\sigma^2R^{4/3}}\right)^{k+2/3}\right)\end{aligned} \tag{33}.$$

In Eq. (33), as in the case of mono-oriented flow, we have assumed the difference between the gradient of the momentum flux and the gradient of the mean value of the same terms computed in the convective prediction and convective correction steps to be negligible.

In Eqq. (31)-(32) the argument of the function *sign*(…) is minus the product of the gradient times the flow rate and the function is equal to 1 or -1 if the argument sign is respectively positive or negative, k is the index of the known time level and variables with index $k + ⅔$ are assumed known in system (33) from the previous solution of systems (31)-(32).

Observe that source terms are neglected in system (32), since the rotational components of the fluxes are assumed to be small with respect to the irrotational ones. Systems (31) is equal to system (8) if the value of function *sign*(…) is equal to 1. In the other case, system (31) degenerates in the ODEs system:

$$\frac{d\sigma}{dt}=0,\quad \frac{dq}{dt}+g\sigma\frac{\partial H^k}{\partial x}+\frac{gn^2q|q|}{\sigma R^{4/3}}=0 \tag{34}.$$

If the value of function *sign*(…) is equal to 1 system (32) degenerates in the steady-state condition:

$$\frac{d\sigma}{dt}=0,\quad \frac{dq}{dt}=0 \tag{35}.$$

After integration in space and time, systems (31)-(33) can be written as:
1) convective prediction system:

$$\left(\sigma_j^{k+1/3}-\sigma_j^k\right)\Delta x+\int_0^{\Delta t}f_{j,j-1}^p dt+\int_0^{\Delta t}f_{j,j+1}^p dt=0 \tag{36,a},$$

$$\left(q_j^{k+1/3}-q_j^k\right)\Delta x+\int_0^{\Delta t}f_{j,j-1}^p ve_{j,j-1}dt+\int_0^{\Delta t}f_{j,j+1}^p ve_{j,j+1}dt=\int_0^{\Delta t}\mathrm{B}_2^p\left(\mathbf{U}_j\right)dt \tag{36,b},$$

$$\text{where} \quad \begin{aligned} f^p_{j,b} &= max\left(0, f_{j,p}\right), \text{ if } \varphi^k_j \geq \varphi^k_b \\ f^p_{j,b} &= min\left(0, f_{j,p}\right), \text{ if } \varphi^k_j < \varphi^k_b \end{aligned} \tag{36,c}$$

2) convective correction system:

$$\left(\sigma_j^{k+2/3} - \sigma_j^{k+1/3}\right)\Delta x + \int_0^{\Delta t} f^c_{j,j-1} dt + \int_0^{\Delta t} f^c_{j,j+1} dt = 0 \tag{37,a}$$

$$\left(q_j^{k+2/3} - q_j^{k+1/3}\right)\Delta x + \int_0^{\Delta t} f^c_{j,j-1} ve_{j,j-1} dt + \int_0^{\Delta t} f^c_{j,j+1} ve_{j,j+1} dt = 0 \tag{37,b}$$

$$\text{where} \quad \begin{aligned} f^c_{j,b} &= min\left(0, f_{j,b}\right), \text{ if } \varphi^k_j \geq \varphi^k_b \\ f^c_{j,b} &= max\left(0, f_{j,b}\right), \text{ if } \varphi^k_j < \varphi^k_b \end{aligned} \tag{37,c}$$

3) diffusive correction system:

$$\left(\sigma_j^{k+1} - \sigma_j^{k+2/3}\right)\Delta x + \int_0^{\Delta t} f_{j,j-1} dt + \int_0^{\Delta t} f_{j,j+1} dt = \left(\overline{f}^{\,p}_{j-1,j} + \overline{f}^{\,p}_{j,j+1} + \overline{f}^{\,c}_{j-1,j} - \overline{f}^{\,c}_{j,j+1}\right)\Delta t \tag{38,a}$$

$$\left(q_j^{k+1} - q_j^{k+2/3}\right)\Delta x + \int_0^{\Delta t} g\left(M_{j,j+1} - M_{j-1,j}\right) dt = \int_0^{\Delta t} \left(g\sigma_j S_0 - \frac{g n^2 q_j^2}{\sigma_j R_j^{4/3}} \right) \Delta x dt - \overline{\mathrm{B}}_2^p \Delta x \Delta t \tag{38,b}$$

where $\overline{f}^{\,p}_{j,j+1}$, $\overline{f}^{\,c}_{j,j+1}$ are the mean values of the volume flux between cells j and j+1, as computed in the convection prediction and correction steps.

Observe that in Eqq. (36) strictly negative (entering) fluxes are always coming from cells with higher potential and strictly positive (leaving) fluxes are always moving to cells with lower potential and viceversa in the convective correction system (37). This allows to solve first the convective prediction step moving from cells with higher to cells with lower potential and the convective correction step moving from cells with lower to cells with higher potential.

The sum of systems (36), (37) and (38) is an approximation of the integral of the original system (1)-(2), and this approximation is as good as the difference between the fluxes and the source terms in the correction systems is either small or time-independent.

3.1 Solution of the convective prediction and convective correction systems

A 1st order spatial approximation of the unknown variables $\sigma(h)$ and q is assumed inside each cell, as in the solution of the convective prediction problem shown in section 2. According to this approximation, as well as to the steady-state assumption of the spatial total head gradient, each system (36)-(37) can be viewed as an ODEs system. Convective prediction system (36) can be written as:

$$\frac{d\sigma_j}{dt}\Delta x + \delta^p f^p_{j,j-1}(t) + \delta^p f^p_{j,j+1}(t) + \left(1-\delta^p\right)\phi^p_{j,j-1} + \left(1-\delta^p\right)\phi^p_{j,j+1} = 0 \tag{39,a}$$

$$\frac{dq_j}{dt}\Delta x + \delta^p f^p_{j,j-1}(t) ve_{j,j-1}(t) + \left(1-\delta^p\right)\xi^p_{j,j-1} + \delta^p f^p_{j,j+1}(t) ve_{j,j+1}(t) + \left(1-\delta^p\right)\xi^p_{j,j+1} = \mathrm{B}_2^p\left(\mathbf{U}_j\right) \tag{39,b}$$

where

$$\delta^p = sign\left(f^p\right) \tag{40}.$$

Convective correction system (37) can be written as:

$$\frac{d\sigma_j}{dt}\Delta x + \delta^c f^c_{j,j-1}(t) + \delta^c f^c_{j,j+1}(t) + \left(1-\delta^c\right)\phi^c_{j,j-1} + \left(1-\delta^c\right)\phi^c_{j,j+1} = 0 \tag{41,a},$$

$$\frac{dq_j}{dt}\Delta x + \delta^c f^c_{j,j-1}(t)ve_{j,j-1}(t) + \left(1-\delta^c\right)\xi^c_{j,j-1} + \delta^c f^c_{j,j+1}(t)ve_{j,j+1}(t) + \left(1-\delta^c\right)\xi^c_{j,j+1} = 0 \tag{41,b},$$

where

$$\delta^c = sign\left(f^c\right) \tag{42}.$$

Systems (39) can be solved sequentially in each computational cell, as previously seen for the case of mono-oriented flow, moving from the higher to the lower approximated potential. After solution of all systems (39), systems (41) can be solved sequentially moving viceversa from the lower to the higher approximated potentials. Of course if an exact scalar potential exists and it is known, the convective correction system reduces to the identity:

$$\sigma\left(h_j\right)^{k+2/3} = \sigma\left(h_j\right)^{k+1/3}, \quad q_j^{k+2/3} = q_j^{k+1/3} \tag{43}.$$

In the convective prediction system, the piezometric gradient of cell *j* at the interface with cell *b* downstream along the flow direction at the beginning of each time step can be computed as in Eq. (19).

The polynomial time approximation of the leaving volume and momentum fluxes are the entering values for the next element in both convective prediction and correction systems. These polynomials can be estimated as function of the $\sigma(h)$ and q values computed during the running time step, as already seen for the case of mono-oriented flow.

3.2 Solution of the diffusive correction system

Diffusive correction system (33) is further simplified and linearized as previously explained for the case of mono-oriented flow, to get the following fully implicit discretization:

$$T_j^{k+2/3}\frac{H_j^{k+1}-H_j^{k+2/3}}{\Delta t}\Delta x + f^{k+1}_{j,j-1} + f^{k+1}_{j,j+1} - f^{k+2/3}_{j,j-1} - f^{k+2/3}_{j,j+1} = 0\,, \quad j=1,\ldots,N \tag{44,a},$$

$$\frac{q_j^{k+1}-q_j^{k+2/3}}{\Delta t} + \left(gn^2 q_j^{k+1}\left(\frac{\left|q_j\right|}{\sigma_j R_j^{4/3}}\right)^{k+\frac{2}{3}} + g\sigma_j^{k+\frac{2}{3}}\frac{\partial H_j^{k+1}}{\partial x} - gn^2\left(\frac{q_j\left|q_j\right|}{\sigma_j R_j^{4/3}}\right)^{k+\frac{2}{3}} - g\sigma_j^{k+\frac{2}{3}}\frac{\partial H_j^{k}}{\partial x}\right) = 0 \tag{44,b};$$

The error due to the linearization of Eqq. (33) goes to zero along with the size of the time step, as well as the difference between the predicted and the corrected fluxes and water levels. The mean value of the known water flux of Eq. (33,a), as well as of the source term of Eq. (33,b) have been approximated in Eqq. (44) at time level $k + ⅔$, that represents the best known time level estimation for the vector variables **U**.

Eq. (44,b) can be written in the form:

$$q_j^{k+1} = -disp_j \frac{\partial H_j^{k+1}}{\partial x} + q_j^{k+2/3} + disp_j \frac{\partial H_j^{k+2/3}}{\partial x} \quad (45),$$

where $disp_j = \dfrac{g\sigma_j^{k+2/3}}{\left(\dfrac{1}{\Delta t} + gn^2 \left(\dfrac{|q_j|}{\sigma_j R_j^{4/3}}\right)^{k+2/3}\right)}$.

Merging Eqq. (45) in Eqq. (29) and these ones in Eq. (44,a), we get:

$$\frac{H_j^{k+1} - H_j^{k+2/3}}{\Delta t}\Delta x + \frac{1}{T_j^{k+2/3}}\left(ds_{j,j-1}^{k+1} + ds_{j,j+1}^{k+1}\right) = \frac{1}{T_j^{k+2/3}}\left(ds_{j,j-1}^{k+2/3} + ds_{j,j+1}^{k+2/3}\right) \quad , \quad j = 1,\ldots,N \quad (46),$$

where

$$ds_{j,b} = -disp_j \frac{\partial H_j}{\partial x} \text{ if } f_{j,b}^k \geq 0 \; ; \quad ds_{j,b} = disp_b \frac{\partial H_b}{\partial x} \text{ if } f_{j,b}^k < 0 \quad (47).$$

After discretization of the total head gradient as given by Eq. (19), Eqq. (46) form a linear algebraic system, with a matrix that is symmetric, positive defined and very well conditioned.

The flow rate unknowns, at time level k+1, can be computed by means of Eq. (45).

3.3 Boundary conditions

Boundary conditions are the same summarized in section 2.3. An approximated potential gradient consistent with the flux sign at time t^k is assumed at the two boundaries of the domain; this implies that zero entering volume and momentum fluxes are assigned at both boundary cells for the solution of convective correction systems (37).

Boundary conditions of the diffusive problem (40) can be obtained by Eqq. (28), changing Eqq. (28,a) and (28,c) respectively with:

$$q_b^{k+1} = q_b^{k+2/3} \quad , \quad q_b^{k+1} = \sigma_b \sqrt{g h_b^{k+2/3}} \quad (48).$$

4 1D NUMERICAL TESTS

4.1 1 Dam-break with finite downstream water depth

In this test the channel is frictionless, horizontal, infinitely large and 2000 m long. Bore is located at x = 1000 m, with a water depth equal to 100 m in the upstream part of the channel and a water depth equal to 1 m in the downstream one. Numerical results have been compared with the analytical ones obtained by *Stoker* (1957), in terms of water depths and flow velocities. Numerical solution is computed at time 9.9 s, using a cell size Δx = 10 m and a time step size Δt = 0.1 s. The maximum value of the CFL number is approximately 0.5. A comparison of the results obtained using 11 different explicit numerical models can be found in *Zoppou and Roberts* (1999). The Authors compared the L_1 norms of errors of water depth and flow velocity defined as:

$$L_{1,h} = \frac{\sum_{j=1,N} \left| h_{ex,j} - h_{nu,j} \right|}{\sum_{j=1,N} \left| h_{ex,j} \right|}, \qquad L_{1,ve} = \frac{\sum_{j=1,N} \left| ve_{ex,j} - ve_{nu,j} \right|}{\sum_{j=1,N} \left| ve_{ex,j} \right|} \tag{49},$$

where sub-indexes *ex* and *nu* indicate respectively the exact solution and the numerical results. The MAST model provides $L_{1,h}$ = 1.32 10^{-2} and $L_{1,ve}$ = 7.63 10^{-2}, where the best norm provided by the 1st spatial approximation order methods is 1.431 10^{-2} for the water depths and 5.5 10^{-2} for the velocities. The worse ranking in the velocity estimation can be explained by the use of the flow rate instead of the flow velocity as unknown, as done in some of the other methods. In figure 4 the comparison between analytical and numerical results is shown, where "sp. step" stands for Δx.

If we assume the computational cost to be proportional to the number of elements times the time steps number, it is possible to maintain the same computational cost by halving the element size and doubling the time step size. In this case, even with a maximum CFL number equal to 3.13, a better performance is obtained for the norm of both the head and the velocity error ($L_{1,h}$ = 7.42 10^{-3} and $L_{1,ve}$ = 4.94 10^{-2}). Also for this case, comparison with the analytical solution is shown in figure 4.

Observe also that for both computational meshes no entropy glitches appear in the numerical solutions. These discontinuities plague several numerical solutions, also obtained with numerical methods with spatial approximation order greater than one (Zoppou and Roberts, 2003).

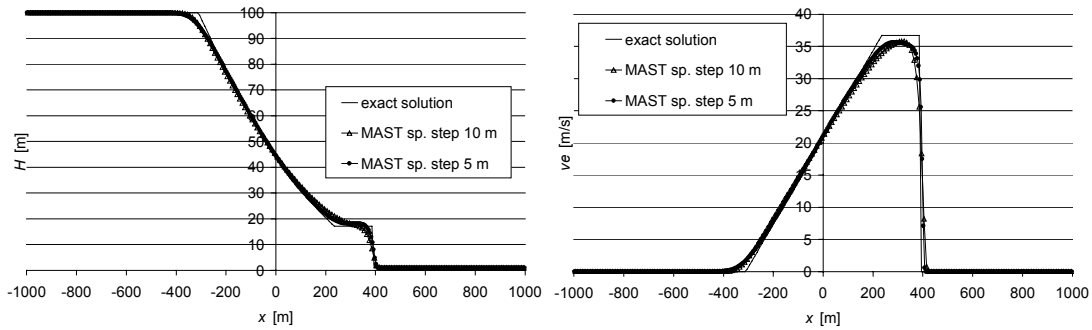

Figure 4. Dam-break with finite downstream water depth. Water levels and flow velocities

4.2 Steady flow over a bump with hydraulic jump

A steady-state transcritical flow over a bump, with a smooth transition followed by a hydraulic jump is simulated. The channel is infinitely large, horizontal, frictionless, 25 m long. Boundary conditions are given by:

$$\begin{aligned} q(0,t) &= 0.18\ \mathrm{m}^3/\mathrm{s} \\ h(0,t) &= 0.33\ \mathrm{m} \end{aligned} \tag{50}.$$

Initial condition is given by a constant water level over the channel, equal to 0.33 m, and bed profile is given by:

$$z(x) = \begin{cases} 0.2 - 0.05(x-10)^2 & \text{if } 8 \le x \le 12 \\ 0 & \text{otherwise} \end{cases} \tag{51}.$$

The test has been carried on using Δx = 0.125 m and Δt = 0.1 s, with a maximum CFL number equal to 3.62. In figure 5 the computed water level and flow rates are compared with the corresponding

analytical solution after 1000 iterations (T = 100 s). Observe that the hydraulic jump is captured in few cells and that no oscillations occur in the water surface and, most important, in the flow rate profile along the bump. Increasing the computational time, flow rate value tends asymptotically to the stationary value q = 0.18 cm/s in all the domain.

For the same test, we also provide in figure 6 a comparison of the convergence history of the proposed scheme with respect to the SGM method by *Zhou et al*. (2001) and to the well-balanced high order WENO scheme by *Vukovic and Sopta* (2002). The global relative error R is defined as (see *Zhou et al*., 2001):

$$R = \sqrt{\sum_{i=1,N} \left(\frac{f_i^n - f_i^{n-1}}{f_i^n} \right)^2} \qquad (52),$$

where f_i^n is the water depth in the i^{th} cell computed at time level n^{th} and index n-1 indicates the corresponding water depth computed at the previous time level. Despite the spatial approximation order, MAST scheme converges much faster than the other two higher order schemes.

Observe, in figure 6, the relative flow errors obtained in the same test by the MAST algorithm. In this case f_i^n in Eq. (52) is the flow rate computed in the cell i.

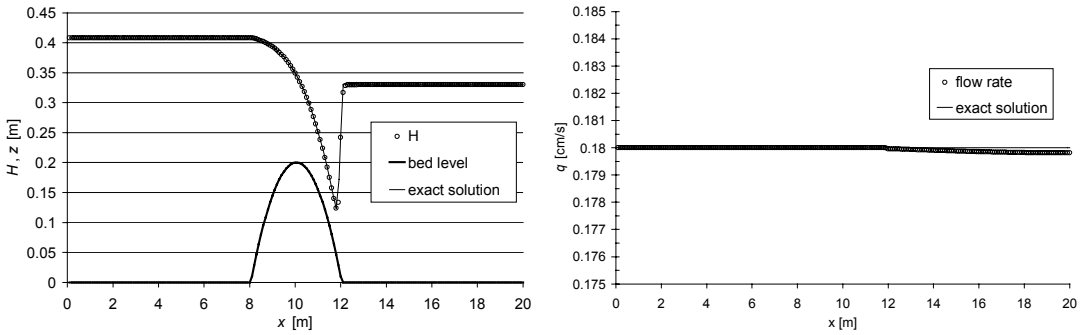

Figure 5. Water surface and flow rate profiles over the bump for transcritical flow after 100 s

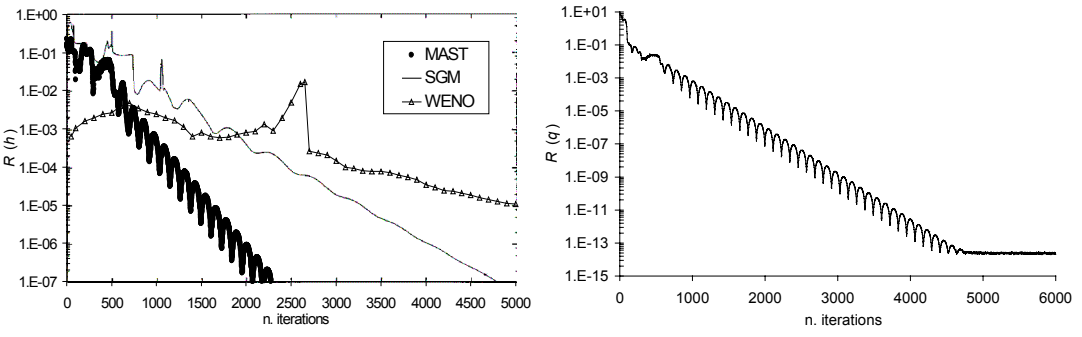

Figure 6. Global relative error of water depths and flow rates for test 4.2

4.3 LeVeque's test problem (1998)

In the LeVeque test problem (*LeVeque*, 1998) a non-stationary flow provided by two dam-breaks propagates in opposite directions; while the upstream moving front passes over a horizontal river bed, the

downstream moving front propagates over a bump. Bed profile is given by:

$$z(x)=\begin{cases}0.25\left(\cos\left(10\pi\left(x-0.5\right)\right)+1\right) \text{ if } 1.4\le x\le 1.6\\ 0 \text{ otherwise}\end{cases} \tag{53}.$$

Channel is assumed frictionless and infinitely wide. The initial conditions are:

$$ve(x,0)=0 \text{ m/s}, \quad h(x,0)=\begin{cases}1+\Delta h-z(x) \text{ if } 1.1\le x\le 1.2\\ 1-z(x) \text{ otherwise}\end{cases} \tag{54},$$

where Δh is the height of the pulse equal to 0.2 m.

In figure 7 computed water levels and flow rates are shown at time t = 0.2 s, using Δx = 0.001 m and Δt = 0.0001 s. The maximum CFL number is 0.355. Observe the discontinuities in both water level and flow rate profiles and also the corresponding errors with respect to the reference solution, computed using a value of the CFL number equal to 3.58 and a number of elements 10 times higher. The error depends on the partial reflection, due to the bump, of the downstream travelling wave in the upstream direction and on the corresponding rapid inversion of the flow direction. The flow inversion is not followed, in the momentum equation of the convective prediction step, by the total head gradient inversion. This error is well corrected by the diffusive correction step, unless the diffusive correction step does not reduce the local change of the unknown variables. This is avoided if the convective change can cross several elements, due to a CFL number much larger than one. Observe, in facts, that refining the mesh using a number of elements 10 times higher oscillations disappear. Similar discontinuities in the numerical solutions have been observed also by *LeVeque* (1998) and by other Authors (*Vukovic S. and Sopta*, 2002; *Zic et al*., 2004).

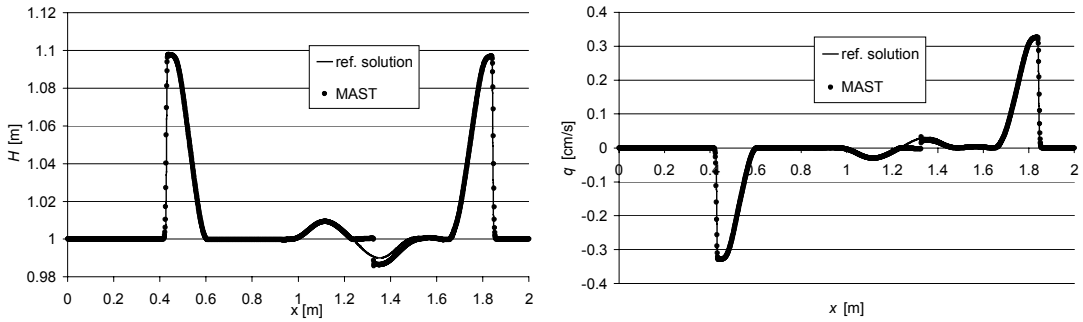

Figure 7. Water level and flow rate for LeVeque's test problem at t = 0.2 s

4.4 Experimental dam-break in a prismatic channel

Some experimental works carried out by the U. S. Army Corps of Engineers (*USACE*, 1960) have been used to compare the model results with experimental data. The experiments have been performed in a 122 m long, 1.22 wide rectangular channel lined with plastic-coated plywood. The bottom slope is 0.5 %, the dam has been placed in the middle of the channel (x = 61 m), and the water depth immediately before the dam is equal to 0.305 m. Two series of experiments have been carried out using both smooth and rough surfaces, with n Manning friction coefficient equal to 0.009 s/m$^{1/3}$ and 0.05 s/m$^{1/3}$ respectively. Initial conditions are given by:

$$h=\begin{cases}0.005x \text{ if } 0\le x\le 61\,\text{m}\\ 0 \text{ if } x>61\,\text{m}\end{cases}, \quad ve(x)=0 \text{ m/s} \tag{55},$$

and are shown in figure 8. At the downstream end of the channel zero water depth is assumed as boundary condition. In the model run Δx = 0.7625 m and Δt = 0.25 s parameters have been used. In figures 9,a-10,b the water depths at abscissas x = 30.5 m, x = 61 m and x = 85.4 m are plotted versus time against the corresponding measured values. Numerical results given by a 2nd order MUSCL scheme (*Zhang et al*., 2003) are also shown. The maximum CFL number is approximately 0.874 for n = 0.009 s/m$^{1/3}$ and 0.82 for n = 0.05 s/m$^{1/3}$. In figure 11 the measured and computed water levels at time t = 10 s for the n Manning coefficient equal to 0.009 s/m$^{1/3}$ are shown. Comparison with experimental data is good and the matching is similar to the one obtained by the 2nd order approximation scheme.

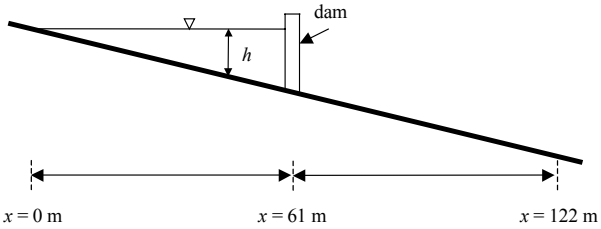

Figure 8. Initial condition for the experimental dam break test

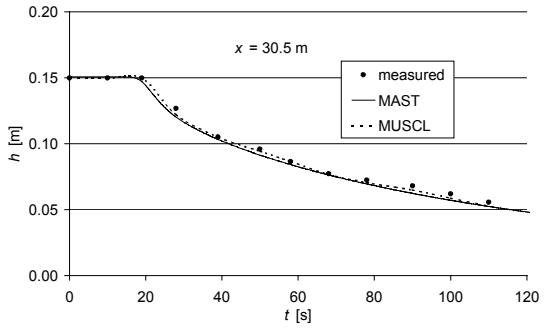

Figure 9,a. Water depths at x = 30.5 m, n = 0.05 s/m$^{1/3}$

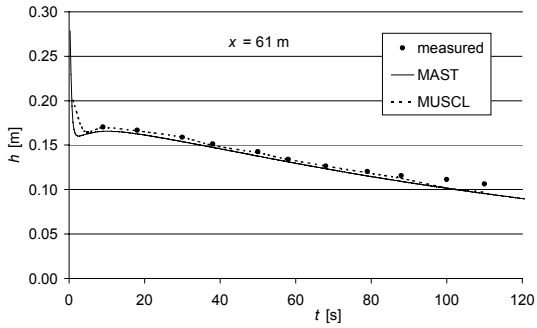

Figure 9,b. Water depths at x = 61 m, n = 0.05 s/m$^{1/3}$

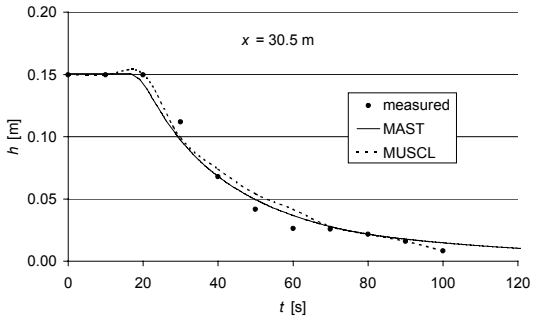

Figure 10,a. Water depths at x = 30.5 m, n = 0.009 s/m$^{1/3}$

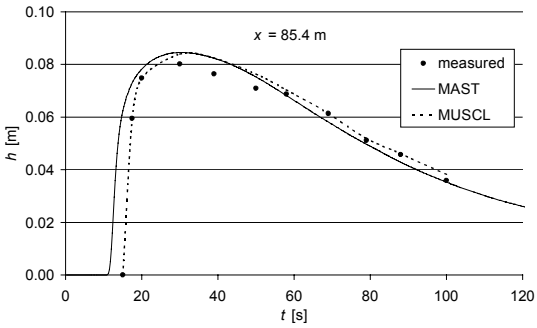

Figure 10,b. Water depths at x = 85.4 m, n = 0.009 s/m$^{1/3}$

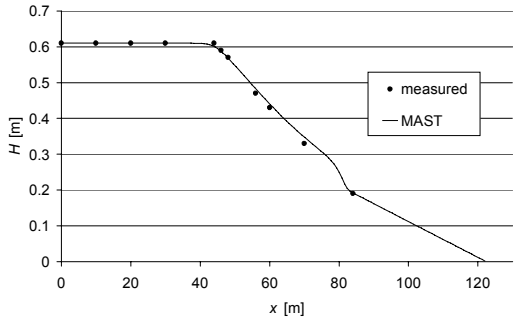

Figure 11. Water level at t = 10 s, n = 0.009 s/m$^{1/3}$

5 THE MAST APPROACH IN THE 2D CASE

5.1 Integral form of the SV equations

The 2D SV equations can be written as:

$$\frac{\partial h}{\partial t}+\frac{\partial uh}{\partial x}+\frac{\partial vh}{\partial y}=0 \tag{56},$$

$$\frac{\partial uh}{\partial t}+\frac{\partial}{\partial x}\left(u^2h\right)+\frac{\partial}{\partial y}\left(uvh\right)+gh\frac{\partial h}{\partial x}+gh\left(\frac{\partial z}{\partial x}+\frac{n^2u\sqrt{\left(uh\right)^2+\left(vh\right)^2}}{h^{7/3}}\right)=0 \tag{57},$$

$$\frac{\partial vh}{\partial t}+\frac{\partial}{\partial y}\left(v^2h\right)+\frac{\partial}{\partial x}uvh+gh\frac{\partial h}{\partial y}+gh\left(\frac{\partial z}{\partial y}+\frac{n^2v\sqrt{\left(uh\right)^2+\left(vh\right)^2}}{h^{7/3}}\right)=0 \tag{58},$$

where x and y are the spatial coordinates, u and v are the x and y velocity components, t is the time, g is the gravitational acceleration, h is the water depth, z is the ground level and n is the Manning friction coefficient. The unknowns in system (56)-(58) are the water depth h and the two flow rates components per unit width uh and vh.

Assume a triangular mesh and a first order approximation of the variables h, uh and vh inside each element. Integration in space of Eqq. (56), (57) and (58), as well as application of the Green's theorem provides:

$$\frac{\partial h}{\partial t}\sigma_e+\sum_{j=1,3}F_{e,j}=0 \tag{59},$$

$$\frac{\partial uh}{\partial t}\sigma_e+\sum_{j=1,3}M^x_{e,j}+R^x_e=0 \tag{60},$$

$$\frac{\partial vh}{\partial t}\sigma_e + \sum_{j=1,3} M^y_{e,j} + R^y_e = 0 \tag{61},$$

where σ_e is the element area, $F_{e,j}$ is the volume flux through the j^{th} side of element e, $M^x_{e,j}$ and $M^y_{e,j}$ are the x and y components of the momentum fluxes along the j^{th} side of element e and the source terms R^x_e, R^y_e are defined as:

$$R^x_e = \sigma_e g\left(h_e \frac{\partial H_e}{\partial x} + \frac{n^2 (uh)_e \sqrt{(uh)^2_e + (vh)^2_e}}{h_e^{7/3}} \right) \tag{62},$$

$$R^y_e = \sigma_e g\left(h_e \frac{\partial H_e}{\partial y} + \frac{n^2 (vh)_e \sqrt{(uh)^2_e + (vh)^2_e}}{h_e^{7/3}} \right) \tag{63},$$

where H_e and h_e are respectively the element water level and water depth.

5.2 Fractional time step decomposition

Assume a scalar value ϕ^k_e, called approximated potential and further specified, to be known at the beginning of each time step in each computational element. The volume and momentum fluxes through each side are defined as function of the flux of the specific flow rate components of both the elements e and ep sharing the given side. The volume flux through the j^{th} side of element e is equal to:

$$FL_{e,j} = (uh)_e \left(y^e_{jp} - y^e_j\right) - (vh)_e \left(x^e_{jp} - x^e_j\right) \tag{64},$$

where jp is the node of element e following node j in counter-clock wise direction.

The volume flux is defined as:

$$F_{e,j} = FL_{e,j} \quad \text{if} \quad FL_{e,j} > 0 \quad \text{and} \quad FL_{e,j} > \quad FL_{ep,m} \tag{65,a},$$

$$F_{e,j} = -FL_{ep,m} \quad \text{otherwise} \tag{65,b},$$

$$M^x_{e,j} = F_{e,j} u_e \;,\; M^y_{e,j} = F_{e,j} v_e \quad \text{if} \; F_{e,j} = FL_{e,j} \tag{66,a},$$

$$M^x_{e,j} = F_{e,j} u_{ep} \;,\; M^y_{e,j} = F_{e,j} v_{ep} \quad \text{otherwise} \tag{66,b},$$

where ep is the index of the element sharing its m^{th} side with the j^{th} side of element e. Observe that condition $F_{e,j} = -F_{ep,m}$ holds for all the internal sides. If $F_{e,j}$ is the positive (leaving) flux of a boundary side, the condition $F_{e,j} = FL_{e,j}$ holds.

After integration in time, prediction step of system (59)-(62) is defined as:

$$\frac{h^{k+1/3} - h^k}{\Delta t}\sigma_e + \frac{1}{\Delta t} \sum_{j=1,3} \int_{\Delta t} F^p_{e,j} dt = 0 \tag{67},$$

$$\frac{(uh)^{k+1/3}-(uh)^{k}}{\Delta t}\sigma_e+\frac{1}{\Delta t}\left(\sum_{j=1,3}\int_{\Delta t}M_{e,j}^{px}dt+\int_{\Delta t}R_e^{px}dt\right)=0 \tag{68},$$

$$\frac{(vh)^{k+1/3}-(vh)^{k}}{\Delta t}\sigma_e+\frac{1}{\Delta t}\left(\sum_{j=1,3}\int_{\Delta t}M_{e,j}^{py}dt+\int_{\Delta t}R_e^{py}dt\right)=0 \tag{69},$$

where:

$$R_e^{px}=\sigma_e g\left(h_e\frac{\partial H_e^k}{\partial x}+\frac{n^2(uh)_e\sqrt{(uh)_e^2+(vh)_e^2}}{h_e^{7/3}}\right) \tag{70,a},$$

$$R_e^{py}=\sigma_e g\left(h_e\frac{\partial H_e^k}{\partial y}+\frac{n^2(vh)_e\sqrt{(uh)_e^2+(vh)_e^2}}{h_e^{7/3}}\right) \tag{70,b},$$

and

$$F_{e,j}^{p}=max\left(0,F_{e,j}\right),\quad M_{e,j}^{px}=F_{e,j}^{p}u_e,\quad M_{e,j}^{py}=F_{e,j}^{p}v_e\quad \text{if } \phi_e^k\geq\phi_{ep}^k \tag{71,a},$$

$$F_{e,j}^{p}=min\left(0,F_{e,j}\right),\quad M_{e,j}^{px}=F_{e,j}^{p}u_{ep},\quad M_{e,j}^{py}=F_{e,j}^{p}v_{ep}\quad \text{if } \phi_e^k<\phi_{ep}^k \tag{71,b}.$$

The first correction system is defined as:

$$\frac{h^{k+2/3}-h^{k+1/3}}{\Delta t}\sigma_e+\frac{1}{\Delta t}\sum_{j=1,3}\int_{\Delta t}F_{e,j}^{c}dt=0 \tag{72},$$

$$\frac{(uh)^{k+2/3}-(uh)^{k+1/3}}{\Delta t}\sigma_e+\frac{1}{\Delta t}\left(\sum_{j=1,3}\int_{\Delta t}M_{e,j}^{cx}dt\right)=0 \tag{73},$$

$$\frac{(vh)^{k+2/3}-(vh)^{k+1/3}}{\Delta t}\sigma_e+\frac{1}{\Delta t}\left(\sum_{j=1,3}\int_{\Delta t}M_{e,j}^{cy}dt\right)=0 \tag{74}.$$

$$F_{e,j}^{c}=max\left(0,F_{e,j}\right)\quad M_{e,j}^{cx}=F_{e,j}^{c}u_e,\quad M_{e,j}^{cy}=F_{e,j}^{c}v_e\quad \text{if } \phi_e^k<\phi_{ep}^k \tag{75,a},$$

$$F_{e,j}^{c}=min\left(0,F_{e,j}\right)\quad M_{e,j}^{cx}=F_{e,j}^{c}u_{ep},\quad M_{e,j}^{cy}=F_{e,j}^{c}v_{ep}\quad \text{if } \phi_e^k\geq\phi_{ep}^k \tag{75,b}.$$

The second correction system is defined as:

$$\frac{h^{k+1}-h^{k+2/3}}{\Delta t}\sigma_e+\frac{1}{\Delta t}\sum_{j=1,3}\int_{\Delta t}F_{e,j}dt=\sum_{j=1,3}\left(\overline{F}_{e,j}^{p}+\overline{F}_{e,j}^{c}\right) \tag{76},$$

$$\frac{(uh)^{k+1}-(uh)^{k+2/3}}{\Delta t}\sigma_e+\frac{1}{\Delta t}\int_{\Delta t}R_e^x dt=\overline{R}_e^{px} \tag{87},$$

$$\frac{(vh)^{k+1}-(vh)^{k+2/3}}{\Delta t}\sigma_e+\frac{1}{\Delta t}\int_{\Delta t}R_e^y dt=\overline{R}_e^{py} \tag{78},$$

where $\overline{F}_{e,j}^{p}$, $\overline{F}_{e,j}^{c}$, $\overline{R}_e^{px}$, $\overline{R}_e^{py}$ are the mean (in time) flux and source term values, as computed in the prediction and in the first correction step. In Eqq. (77)-(78) we neglect the difference between the mean (in time) convective inertia terms and the same mean values computed from the solution of the prediction plus the first correction systems. This is equivalent to assume, in the second correction system:

$$\frac{1}{\Delta t}\sum_{j=1,3}\int_{\Delta t}M_{e,j}^{x}dt\approx\sum_{j=1,3}\overline{M}_{e,j}^{px}+\sum_{j=1,3}\overline{M}_{e,j}^{cx} \tag{79,a},$$

$$\frac{1}{\Delta t}\sum_{j=1,3}\int_{\Delta t}M_{e,j}^{y}dt\approx\sum_{j=1,3}\overline{M}_{e,j}^{py}+\sum_{j=1,3}\overline{M}_{e,j}^{cy} \tag{79,b}.$$

This implies that convective fluxes are missing in Eqq (86)-(88) and we call the prediction system convective prediction system and the first and the second correction systems respectively convective and diffusive correction systems.

Initial conditions of the convective correction system are the final values of the convective prediction system and initial values of the diffusive correction system are the final values of the convective correction system. Spatial piezometric gradient $\partial H^k/\partial x$ is assumed constant in time in all the systems. Observe that summing all the terms of Eqq. (67)-(88) the integrals of Eqq. (59)-(61) are formally found again.

5.3 Solution of the convective prediction and the convective correction systems

A first order spatial approximation is applied to the unknown variables *h*, *uh* and *vh*. Observe that, according to the flux definitions given in Eqq. (71) and (75), the flux integrals from element *e* to element *ep* in the convective prediction step are only function of the element *e* unknowns if $\phi_e^k\geq\phi_{ep}^k$ and are only function of the element *ep* unknowns if $\phi_e^k<\phi_{ep}^k$. In the convective correction step the opposite holds. This allows to solve each system as a sequence of small ODEs systems, one for each computational cell, after ordering the cells according to the their scalar potential. The priority is given to the cells with higher potential in the convective prediction step and to the cells with lower potential in the convective correction step.

In the prediction case, the ODEs system is:

$$\frac{dh_e}{dt}\sigma_e+\sum_{j=1,3}F_{e,j}^{p}(t)=0 \tag{80},$$

$$\frac{d(uh)_e(t)}{dt}\sigma_e+\sum_{j=1,3}\left(M_{e,j}^{px}(t)+R_e^{px}(t)\right)=0 \tag{81},$$

$$\frac{d(vh)_e(t)}{dt}\sigma_e + \sum_{j=1,3}\left(M^{py}_{e,j}(t) + R^{py}_e(t)\right) = 0 \tag{82}.$$

Call *out* (*in*) the index of any side shared with any element *ep* such that $\phi^k_e \geq \phi^k_{ep}$ in the convective prediction step ($\phi^k_e < \phi^k_{ep}$ in the convective correction step). After the single system of element *e* is solved along the time step, the mean values of the fluxes leaving through each side *out* have to be estimated.

The mean value of the leaving volume flux $\overline{F}^p_{e,out}$ can be estimated by partitioning the total mean leaving flux, obtained from the mass balance in the element and equal to:

$$\sum_{out}\overline{F}^p_{e,out} = -\sum_{in}\overline{F}^p_{e,in} - \frac{h^{k+1/3}_e - h^k_e}{\Delta t}\sigma^e \tag{83},$$

Observe that the mean values of the entering fluxes $\overline{F}^p_{e,in}$ are known from the solution of the elements with higher scalar potential. Partition can be done adopting for each element side a weight equal the arithmetic mean between the initial and the final flux values, that is:

$$\overline{F}^p_{e,j} = \frac{F^p_{e,j}(0) + F^p_{e,j}(\Delta t)}{\sum_{out}\left(F^p_{e,out}(0) + F^p_{e,out}(\Delta t)\right)}\sum_{out}\overline{F}^p_{e,out} \tag{84},$$

where *j* is also the index of any side shared with an element with lower potential.

Given the initial, the mean and the final values, a parabolic volume flux time approximation in each side *j* is finally computed and used to estimate the entering fluxes in the connected elements with lower approximated potential. Conservation of the mean values can be easily proved to guarantee the local and global mass conservation.

The mean leaving momentum fluxes $\overline{M}^{px}_{e,j}, \overline{M}^{py}_{e,j}$ can also be estimated in a similar way. After integration, they can be computed as:

$$\overline{M}^p_{e,j} = \frac{F^p_{e,j}(0) + F^p_{e,j}(\Delta t)}{\sum_{out}\left(F^p_{e,out}(0) + F^p_{e,out}(\Delta t)\right)}\sum_{out}\overline{M}^p_{e,out} \tag{85}.$$

Due to the non-linearity of the momentum balance equations, the mean value of the total leaving momentum flux at the r.h.s. of Eq. (85) has to be numerically estimated from the single time values. In the coded algorithm, the leaving momentum fluxes in three Gauss points, selected in the time interval 0 – Δt have been computed using a C^1 interpolation of the solution values produced by the Runge-Kutta method adopted for the solution of the ODEs system. A parabolic momentum flux estimation is finally carried out for each element side using the initial, the final and the mean value. A similar procedure is carried out for the solution of the convective correction system, written as:

$$\frac{dh_e}{dt}\sigma_e + \sum_{j=1,3}F^c_{e,j}(t) = 0 \tag{86},$$

$$\frac{d(uh)_e(t)}{dt}\sigma_e + \sum_{j=1,3}M^{cx}_{e,j}(t) = 0 \tag{87},$$

$$\frac{d(vh)_e(t)}{dt}\sigma_e + \sum_{j=1,3} M^{py}_{e,j}(t) = 0 \tag{88}.$$

Observe that the source terms have been totally allocated in the convective prediction step. This simplifies the solution of the problem and is computationally ineffective due to the small size of the corrective fluxes, sought after with a good choice of the approximated potential function.

5.4 The approximated scalar potential

Fractional step methodology provides accurate results only if the solution of the convective prediction step is close to the corrected one; otherwise, the computation of the spatial gradients at different time levels can strongly affect the solution. In our case, to minimize the variable change in the convective correction step, it is important to choose an approximated potential with a gradient flux opposite in sign, as much as possible, to the water flux along the element sides. If an exact potential exists and this condition is always attained, convective correction system vanishes in the following identities:

$$h_e^{k+2/3} = h_e^{k+1/3}, \quad (uh)_e^{k+2/3} = (uh)_e^{k+1/3}, \quad (vh)_e^{k+2/3} = (vh)_e^{k+1/3} \tag{89}.$$

We seek to minimize, at the known time level t^k, the following functional:

$$F = \sum_j -min\left(0, (uh)_e\left(y^e_{jp} - y^e_j\right) - (vh)_e\left(x^e_{jp} - x^e_j\right)\right) \tag{90},$$

where j is again the index of any side shared by an element ep such that:

$$\phi^k_e \geq \phi^k_{ep} \tag{91}.$$

The best set of scalar potentials, that minimizes function (90) subject to constraints (91), is the solution of a global minimum search, computationally very expensive. Excellent results can also be found more easily by performing a local search, if a good starting point is chosen. This point is the potential set corresponding to numerically estimated gradients as opposite as possible to the velocity gradients times the water depth. In a 3D space, it can be viewed as the set of potentials that are as close as possible to the planes matching the potential value at the triangle centre, with a gradient opposite to the velocity gradient times the water depth (see figure 12). The functional component corresponding to the side j of each element is weighted with the absolute value of the flux and to all the element components a penalty term related to the change with respect to the previous value is added, to finally get:

$$min(F') = \sum_{e=1}^{N}\sum_{j=1}^{3}\left[\left(\phi^k_{ep} - \phi^k_e + \left(x_{ep} - x_e\right)(uh)^k_e + \left(y_{ep} - y_e\right)(vh)^k_e\right)^2 \left|FL_{e,j}\right| + \alpha\left(\phi^k_e - \phi^{k-1}_e\right)^2\right] \tag{92},$$

where x_e, y_e are the coordinates of the centre of element e, and α is a small positive number.

The proposed functional is convex and the minimum can be found by setting to zero the partial derivatives with respect to all the element approximated potential. The resulting linear system is sparse, symmetric, positive definite and well conditioned. The penalty term in the square brackets is aimed to avoid in the linear system zero diagonal terms when the specific flow rate components at time t^k are equal to zero. Increasing the penalty term, a potential distribution more similar to the previous one computed at time t^{k+1} is obtained. To obtain a compromise between computational efficiency and the approximated potential quality, the coefficient α of the penalty term can be normalized with respect to the diagonal term of the corresponding original equation, with a lower boundary given by the machine precision.

After the minimum of F' is found, the approximated potentials can be further refined with a local search, iteratively adjusting the element potential values until a local minimum of the functional F is found.

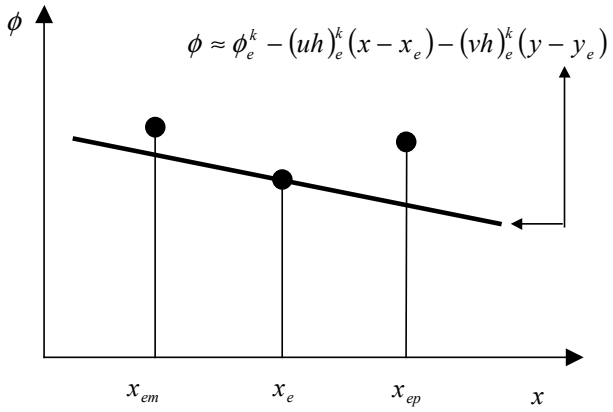

Figure 12. Linear approximation of the approximated potential around a central value

5.5 Computation of the gradients for the convective systems

Water level gradient is usually estimated, in the case of triangular mesh and first order approximation, by applying Green's theorem at the 2D integral of the water level over the element (see for example *Anastasiou and Chan*, 1997; *Putti et al*., 1990). This estimation is accurate only if regular meshes are used. In the case of unstructured, automatically generated meshes, it is affected by large errors and heterogeneous gradients are obtained also in the case of a linear variation of the water level inside the domain. A most robust technique has been developed, similar to the algorithm proposed by *Hubbard* (1999).

The gradient is computed as a weighted average between one or two vectors, each one corresponding to a vertex *j* of the triangle. Each j^{th} vector $\nabla_{ej}\mathbf{H}$ is computed assuming a linear variation of the water level inside a new triangle, defined by the centre of element *e* and the centres of the two surrounding elements sharing the j^{th} node. Zero weight is given to the j^{th} vector corresponding to a negative scalar products between the element velocity and the direction of the line connecting the centre of element *e* with its j^{th} vertex. Each vector is the gradient of the linear function matching the water levels in the centre of element *e* and in the centre of elements *ep*, *em* sharing vertex *j* with triangle *e* (see figure 13). The final gradient is given by:

$$\nabla\mathbf{H} = \frac{\sum_{j=1}^{3} max\left(0, \mathbf{n}_{ej}\mathbf{n}_{uv}\right)\nabla_{ej}\mathbf{H}}{\sum_{j=1}^{3} max\left(0, \mathbf{n}_{ej}\mathbf{n}_{uv}\right)} \quad (93),$$

where $\mathbf{n}_{ej}\,\mathbf{n}_{uv}$ is the scalar product between the centre-vertex and the velocity directions.

If element *e* has an edge on the impervious boundary, its velocity is parallel to the boundary and one of the three conditions required to compute the gradient corresponding to the downstream vertex is missing. This condition is replaced by the equality of the expected total piezometric gradient direction and the impervious edge direction.

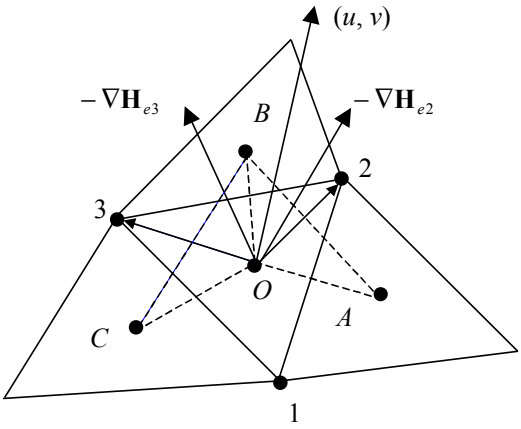

Figure 13. Computation of the spatial water level gradient

5.6 Boundary conditions for the convective prediction system

At the beginning of each time steps, elements with one boundary side j are selected. The Froude number $fr_{e,j}$ of the flux per unit length is computed as:

$$fr_{e,j} = \frac{FL_{e,j}}{L_{e,j}\, h_e^{3/2}\sqrt{g}} \quad \text{if } FL_{e,j} \geq 0, \qquad fr_{e,j} = \frac{h_e^b u_e^b \left(y_{jp}^e - y_j^e\right) - h_e^b v_e^b \left(x_{jp}^e - x_j^e\right)}{L_{e,j}\, h_e^{3/2}\sqrt{g}} \quad \text{if } FL_{e,j} < 0 \tag{94},$$

where $L_{e,j}$ is the length of the j^{th} side of element e. Boundary conditions of the prediction problem (67)-(69) depend on the flux sign and on the Froude number $fr_{e,j}$. One of the following cases can be selected:
1) Flux through the boundary side is entering the domain. In this case the following equalities hold:

$$F_{e,j}^p = BF_{e,j}, \quad M_{e,j}^{px} = BM_{e,j}^x, \quad M_{e,j}^{py} = BM_{e,j}^y \tag{95},$$

$$\begin{aligned} BF_{e,j} &= h_e^b u_e^b \left(y_{jp}^e - y_j^e\right) - h_e^b v_e^b \left(x_{jp}^e - x_j^e\right), \\ BM_{e,j}^x &= BF_{e,j} u_e^k, \quad M_{e,j}^{py} = BF_{e,j} v_e^k \end{aligned} \quad \text{if } fr_{e,j} \leq 1 \tag{96},$$

$$\begin{aligned} BF_{e,j} &= h_e^b u_e \left(y_{jp}^e - y_j^e\right) - h_e^b v_e \left(x_{jp}^e - x_j^e\right), \\ BM_{e,j}^x &= BF_{e,j} u_e, \qquad BM_{e,j}^y = BF_{e,j} v_e \end{aligned} \quad \text{if } fr_{e,j} > 1 \tag{97},$$

where h_e^b, u_e^b and v_e^b are the assigned boundary water depths and velocity components;
2) Flux through the boundary side is leaving the domain. In this case water depth gradient is not known and convective prediction system is replaced, for element e, with the hypothesis of neglecting the flow rate and momentum local changes in the same elements. This is equivalent to say that the leaving volume and momentum fluxes are equal to the entering ones. From the previous hypothesis, we obtain:

$$h_e^{k+1/3} = h_e^k, \quad u_e^{k+1/3} = \frac{\sum_{in} M_{e,in}^x(0)}{\sum_{in} F_{e,in}(0)}, \quad v_e^{k+1/3} = \frac{\sum_{in} M_{e,in}^y(0)}{\sum_{in} F_{e,in}(0)} \tag{98},$$

with the same meanings specified for Eqq. (65)-(66).
3) Flux through the boundary side is zero (impervious boundary). In this case the following conditions hold:

$$F_{e,j}^p = 0, \quad \nabla \mathbf{H}_e \cdot \mathbf{n}_{e,j} = 0 \tag{99},$$

where $\nabla \mathbf{H}_e$ is the water level gradient and $\mathbf{n}_{e,j}$ is the unit vector normal to the impervious side.

We assume the scalar potential immediately outside the boundary sides to be consistent with the flux sign. This implies the volume and momentum fluxes to be zero in the convective correction step, that is:

$$F_{e,j}^c = 0, \quad M_{e,j}^{cx} = 0, \quad M_{e,j}^{cy} = 0 \tag{100}.$$

5.7 Solution of the convective steps with small water depth

Instabilities occur in the solution of the convective steps when initial water depth values are negative, and convergence is very slow when positive, but very small values are computed in the tail of the propagating wave. Small negative water depths can be computed by the diffusive system in the preceding time step, when the result of the convective steps is a very small positive value.

The following procedure is followed in order to avoid these inconveniences and to maintain the local mass conservation property. A very small minimum water depth h_{min} is chosen (10^{-4}-10^{-5} m) and both the momentum flux and the friction terms are computed, for $h < h_{min}$, as the first order expansion around h_{min}, that is:

$$R_e^{px} = \sigma_e g \left[h_e \frac{\partial H_e^k}{\partial x} + \frac{n^2 (uh)_e \sqrt{(uh)_e^2 + (vh)_e^2}}{h_{min}^{7/3}} - (h_e - h_{min}) \frac{7}{3} \frac{n^2 (uh)_e \sqrt{(uh)_e^2 + (vh)_e^2}}{h_{min}^{10/3}} \right]$$
$$R_e^{py} = \sigma_e g \left[h_e \frac{\partial H_e^k}{\partial y} + \frac{n^2 (vh)_e \sqrt{(uh)_e^2 + (vh)_e^2}}{h_{min}^{7/3}} - (h_e - h_{min}) \frac{7}{3} \frac{n^2 (vh)_e \sqrt{(uh)_e^2 + (vh)_e^2}}{h_{min}^{10/3}} \right] \tag{101},$$

$$M_{e,j}^{px} = F_{e,j}^p \frac{(uh)_e}{h_{min}} - (h_e - h_{min}) F_{e,j}^p \frac{(uh)_e}{h_{min}^2}, \quad M_{e,j}^{py} = F_{e,j}^p \frac{(vh)_e}{h_{min}} - (h_e - h_{min}) F_{e,j}^p \frac{(vh)_e}{h_{min}^2} \tag{102}.$$

This change allows a fast solution of the problem for small, but positive water depth values. When a tail occurs, solution converges anyway to negative values. To avoid this, a control is set in the ODEs solver in order to return the time value corresponding to a zero water depth value. When the solver stops before the end of the time step, the remaining water volume entering the cell is left in the same cell assuming zero leaving flux after the return time. When the initial water depth h_e^k (or $h_e^{k+1/3}$ in the convective correction step) is negative, two possibilities occur. Call *Vol* the volume corresponding to the cell area and to the opposite of the negative initial water depth. If the entering volume $V_{in} = -\sum_{in} \overline{F}_{e,in}^p \Delta t$ is smaller than *Vol*, increase h_e^k, assume a zero leaving flux, and move to the next cell. If V_{in} is greater then *Vol*, set $h_e^k = 0$, reduce the mean entering volume and momentum fluxes according to the $Vol / \Delta t$

value and solve the ODEs system.

5.8 Solution of the diffusive correction system

Diffusive correction system (76)-(78) is simplified by changing the integral mean values with the fully implicit time discretization and linearizing the source terms. This provides a diffusive effect, that goes to zero along with the size of the source terms and the diffusive fluxes. This size is proportional to the time step, as can be observed in Eqq. (77)-(78). According to these hypotheses, diffusive correction system (76)-(78) is equivalent to the following fully implicit time discretization:

$$\frac{H_e^{k+1} - H_e^{k+2/3}}{\Delta t} + \frac{\partial (uh)^{k+1}}{\partial x} + \frac{\partial (vh)^{k+1}}{\partial y} = 0 \tag{103},$$

$$\frac{(uh)_e^{k+1} - (uh)_e^{k+2/3}}{\Delta t} + g h_e^{k+2/3} \frac{\partial H_e}{\partial x} + \frac{g\, n^2 (uh)_e^{k+1} \left(\sqrt{(uh)_e^2 + (vh)_e^2} \right)^{k+2/3}}{\left(h_e^{7/3} \right)^{k+2/3}} = g h_e^{k+2/3} \frac{\partial H_e^k}{\partial x} + \frac{g\, n^2 \left((uh)_e \sqrt{(uh)_e^2 + (vh)_e^2} \right)^{k+2/3}}{\left(h_e^{7/3} \right)^{k+2/3}} \tag{104},$$

$$\frac{(vh)_e^{k+1} - (vh)_e^{k+2/3}}{\Delta t} + g h_e^{k+2/3} \frac{\partial H_e}{\partial y} + \frac{g\, n^2 (vh)_e^{k+1} \left(\sqrt{(uh)_e^2 + (vh)_e^2} \right)^{k+2/3}}{\left(h_e^{7/3} \right)^{k+2/3}} = g h_e^{k+2/3} \frac{\partial H_e^k}{\partial y} + \frac{g\, n^2 \left((vh)_e \sqrt{(uh)_e^2 + (vh)_e^2} \right)^{k+2/3}}{\left(h_e^{7/3} \right)^{k+2/3}} \tag{105},$$

that is:

$$(uh)_e^{k+1} = -elem_e \frac{\partial H_e^{k+1}}{\partial x} + cos\, t_e^x + (uh)_e^{k+2/3} \tag{106},$$

$$(vh)_e^{k+1} = -elem_e \frac{\partial H_e^{k+1}}{\partial y} + cos\, t_e^y + (vh)_e^{k+2/3} \tag{107},$$

where:

$$elem_e = \frac{g \left(h_e^{10/3} \right)^{k+2/3} \Delta t}{\left(h_e^{7/3} \right)^{k+2/3} + \Delta t\, g\, n^2 \left(\sqrt{(uh)_e^2 + (vh)_e^2} \right)^{k+2/3}} \tag{108,a},$$

and $\quad cos\, t_e^x = elem_e \frac{\partial H_e^k}{\partial x}, \quad cos\, t_e^y = elem_e \frac{\partial H_e^k}{\partial y}$ (108,b),

where the variables with index $k + ⅔$ are known from the solution of the convective correction system.

Merging Eqq. (106)-(108) in Eq. (103), the following relationship is obtained:

$$\frac{H_e^{k+1} - H_e^{k+2/3}}{\Delta t}\sigma_e + \sum_{j=1,3} \delta_{j,jp}^e F_{e,j}^{k+1} = 0 \tag{109},$$

where diffusive fluxes are discretized as:

$$F_{e,j}^{k+1} = -elem_e \frac{H_{ep}^{k+1} - H_e^{k+1}}{\sigma_e} L_{e,j}^2 + \\ cos\, t_e^x \left(y_{jp}^e - y_j^e\right) + cos\, t_e^y \left(x_{jp}^e - x_j^e\right) \quad \text{if } FL_{e,j}^k \geq 0 \tag{110},$$

$$F_{e,j}^{k+1} = -elem_{ep} \frac{H_{ep}^{k+1} - H_e^{k+1}}{\sigma_{ep}} L_{e,j}^2 + \\ cos\, t_{ep}^x \left(y_{jp}^e - y_j^e\right) + cos\, t_{ep}^y \left(x_{jp}^e - x_j^e\right) \quad \text{if } FL_{e,j}^k < 0 \tag{111},$$

$\delta_{j,jp}^e = 1$ if the j^{th} and jp^{th} nodes of element e are shared by the elements e and ep (internal side), $\delta_{j,jp}^e = 0$ if not (boundary side). After computation of the new water levels H_e^{k+1}, the flow rates per unit width can be computed with Eqq. (106)-(108) using the new convective gradients ($\frac{\partial H_e^{k+1}}{\partial x}, \frac{\partial H_e^{k+1}}{\partial y}$) computed as explained in section 5.5.

5.9 Boundary conditions of the diffusive correction system

If $fr_{e,j} > 1$, zero diffusive flux is assigned, that is $F_{e,j}^{k+1} = 0$. If $fr_{e,j} \leq 1$ two possibilities exist. If the assigned boundary water depth is smaller than the critical depth corresponding to the specific flow rate, that is:

$$h_e^b \leq \left(\frac{(uh)_e^2 + (vh)_e^2}{g} \right)^{1/3} \tag{112},$$

a total flux corresponding to the critical depth is assigned to the element boundary side, that is:

$$F_{e,j}^{k+1} = \sqrt{g} \left(h_e^{k+2/3}\right)^{3/2} L_{e,j} - F_{e,j}^{k+2/3} \tag{113}.$$

If inequality (112) does not hold, boundary water depth is assigned to element e, along with the corresponding Dirichlet condition and diffusive flux is computed a posteriori by means of Eq. (109).

6 2D NUMERICAL TESTS

In the following, results of the MAST model are compared with experimental data of two laboratory tests and with results obtained for the simulation of the same experiments by other authors. In all the simulations, the computational mesh is unstructured and automatically designed by the ARGUS ONE mesh generator (*Argus Holdings*). The water depth and velocity components in a given point with x, y coordinates have been assumed equal to the h_e, u_e and v_e values of the element containing the given point, without any further interpolation.

6.1 2D dam-break experiment by Fraccarollo and Toro (1995)

The experimental flume, shown in figure 14, is a rectangular bottom plane partially occupied by a reservoir. The whole bottom plane may rotate around a hinge. This allows to select any constant value of the bottom slope. The experimental flume is 2 m wide and 3 m long, with a vertical wall at 1/3 of its length. The downstream part of the plane bottom is initially dry. Both the walls and the bottom are covered with the same material, corresponding to a Manning coefficient of 0.0095 $s/m^{1/3}$. The width of the movable gate, symmetrically centred, is 0.40 wide m. The flood-plain boundaries are all open.

The Authors measured the pressure at the bottom of the channel by using pressure gauges, the water depths by means of wave height meters and the velocity components by means of an electromagnetic velocity meter. Measure points are reported in figure 14 and their spatial coordinates are reported in table 1.

Two sets of runs have been carried out: the first one assuming zero bottom slope in x and y directions, the second one assuming zero bottom slope in y direction and 7 per cent in x direction.

Numerical results have been compared with the experimental data, with the results of F*raccarollo and Toro* (1995) and with the results of by *Gottardi and Venutelli* (2004).

Fraccarollo and Toro (1995) applied a WAF (Weighted Averaged Flux) scheme (*Toro*, 1989), a 2nd order conservative, shock-capturing Godunov-type scheme. The Authors discretized the domain by means of a regular mesh of 150 × 50 points along x and y directions respectively, with a corresponding element area equal to $8.0\ 10^{-4}$ m^2.

Gottardi and Venutelli (2004) applied an explicit 2nd order central scheme, initially proposed by *Kurganov and Tadmor* (2000); integration in time has been performed by means of a 3rd order TVD-Runge-Kutta scheme. They discretized the domain with regular quadrilateral elements spacing $\Delta x = \Delta y =$ 0.02 m. For MAST simulations, spatial domain is discretized using an unstructured triangular mesh (see figure 15). The total number of elements is 8650, with a mean area approximately equal to $7.0\ 10^{-4}$ m^2.

In the first experiment with zero bottom slope, an initial water depth of 0.6 m in the reservoir has been assigned. In figure 16 the contour plot of the water depths obtained with the MAST (right) and with the Gottardi and Venutelli (K-T) models (left) at the simulation time $T = 0.5$ s are shown. The maximum CFL number obtained with the K-T model is 0.5; in the MAST model a time step 0.01 s has bees used, with a corresponding maximum CFL number approximately equal to 3.62.

The asymmetry of the MAST, more evident for the smaller water depths, is due to the mesh asymmetry. In a previous preliminary work (*Aricò et al*., 2006), where a regular equilateral triangular mesh has been used, perfectly symmetric results can be observed.

In figures 17,a-17,d computed water depths at points “O”, “U1”, “U2” and “D” are compared up to the simulation time 10 s with the corresponding measured values. In figures 17,a-17,c the numerical results by *Fraccarollo and Toro* (1995) are shown for the points “O”, “U1”, “U2”. The MAST results show some diffusive effect missing in the WAF results, because of the higher approximation order with respect to proposed one.

MAST and WAF results are in good agreement with the experimental ones at points “U1”, “U2” and “D”. Observe the difference between measured and numerical results for small times at point “O”, where the shallow water hypothesis does not hold. After the sudden opening of the gate a strong rarefaction wave starts moving in upstream direction. As in the 1D case, water depth at the gate location (point “O” in the specific case) drops to a local minimum value (about 4/9 of the initial depth) after the opening of the gate. This is analogous to the 1D solution by *Ritter* (1992) in a horizontal and frictionless channel. After the minimum is reached, a rising stage follows, due to the fluxes coming from the wall boundaries and to 2D effects.

The delay between measured and computed data at point “O” is likely due to the small time required for the real opening of the gate (about 0.1 s).

Inside the reservoir, *Fraccarollo and Toro* (1995) measured also the static pressure values at the bottom. Pressure measures are reported in meters of water column. For measurement points “U1” and “U2” pressures and water levels values are very close to each other and only water levels are reported in figures 17,b and 17,c; at point “O”, as expected, measured levels and hydrostatic pressures do not match because of the vertical velocity components.

In the second experiment a 7 per cent slope along x direction has been assigned to the bottom plane and the initial water depth value, measured at the wall foundation, is 0.64 m.

In figure 18 the contour plot of the water depths are shown for the proposed model (right) and for the K-T model (left) at the simulation time T = 0.5 s. Maximum CFL number obtained in the run of the K-T model is again 0.5. In the MAST computation Δt =0.01 s has been used, with a corresponding maximum CFL number almost equal to 4.22. Observe also in this case the asymmetry of the MAST results, similar to the asymmetry obtained in the case of zero bottom slope.

In figures 19,a-19,b computed water depths at points "O" and "U1" are compared with the corresponding measured values, as well as with the results of *Fraccarollo and Toro* (1995) up to the simulation time T = 10 s. Observations similar to the previous ones, drawn for the zero bottom slope experiment, can be done.

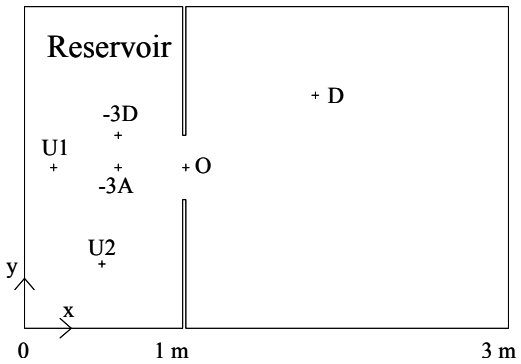

Figure 14. Experimental flume of Fraccarollo and Toro [30]

Point	x [m]	y [m]
U1	0.18	1
U2	0.48	0.4
O	1	1
D	1.802	1.45

Table 1. Spatial coordinates of the experimental points for test 3.1

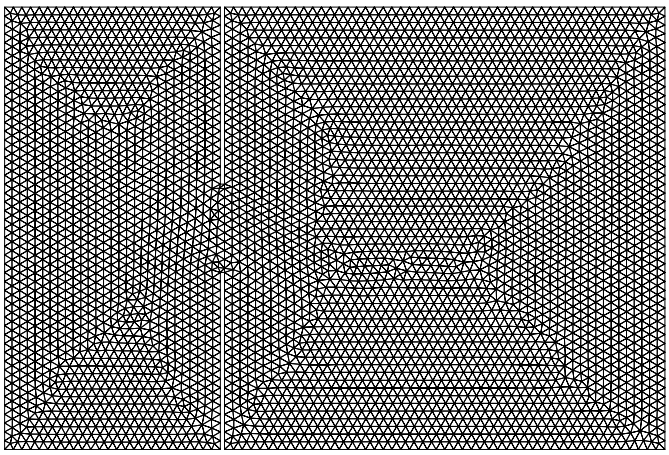

Figure 15. Mesh used for MAST numerical simulation of test 6.1

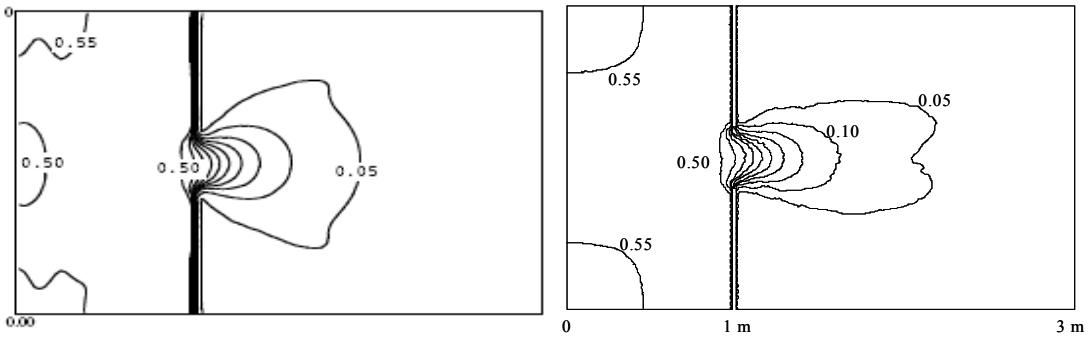

Figure 16. Contour plot of the water depths after 0.5 s by Gottardi and Venutelli [31] and by MAST scheme (0 bottom slope)

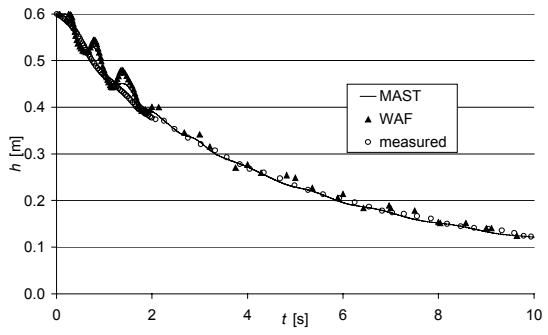

Figure 17,a. Water depths at point "U1" (0 bot. slope)

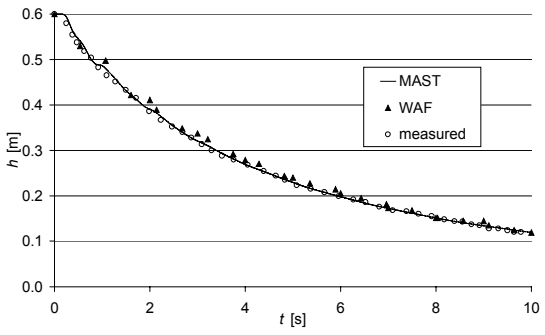

Figure 17,b. Water depths at point "U2" (0 bot. slope)

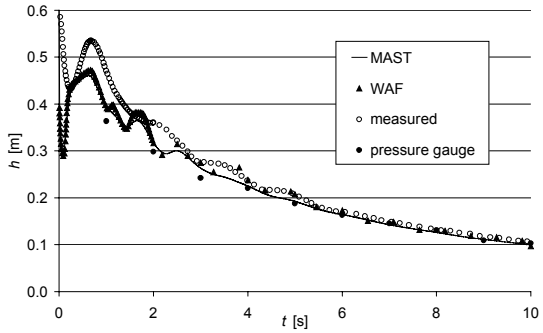

Figure 17,c. Water depths at point "O" (0 bot. slope)

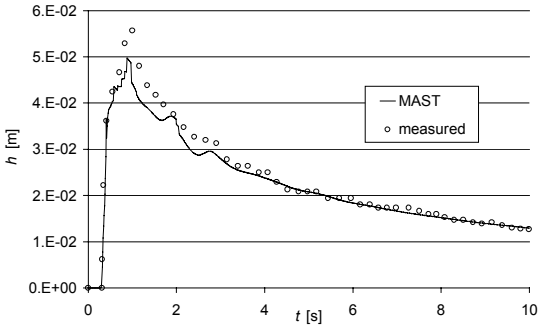

Figure 17,d. Water depths at point "D" (0 bot. slope)

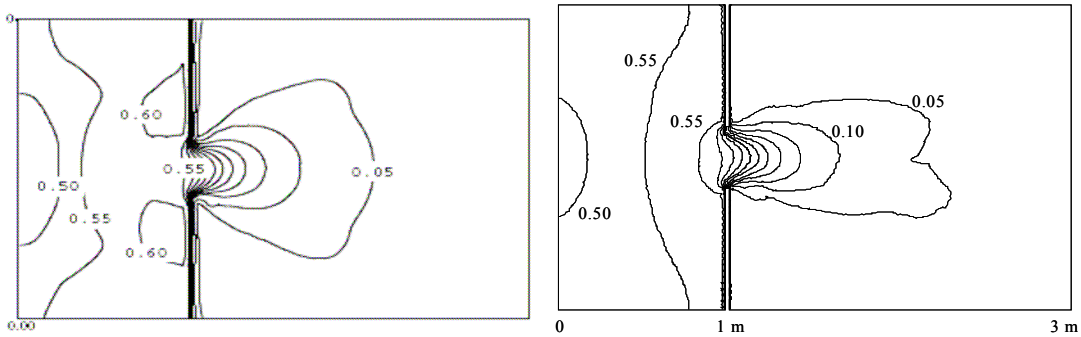

Figure 18. Contour plot of the water depths after 0.5 s by Gottardi and Venutelli [31] and by MAST scheme (0.07 bottom slope)

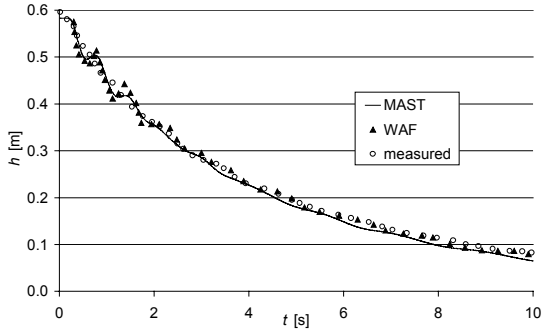

Figure 19.a. Water depths at point "U1" (0.07 bot. slope)

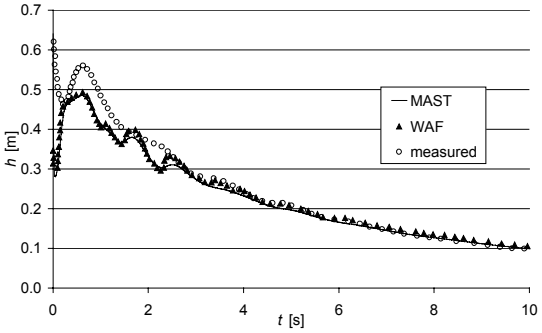

Figure 19,b. Water depths at point "O" (0.07 bot. slope)

6.2 2D dam-break experiment in a L-shaped channel (1999)

Two sets of experiments have been carried out in the Civil Engineering Laboratory of the Catholic University of Louvain (Belgium) (*Soares Frazão* et al., 1999) in a channel with a L-shaped form and rectangular cross section (see figure 20). Bottom slope is zero in both x and y directions.

The upstream reservoir is a tank with rectangular 2.44 m × 2.39 m planar section, closed with a vertically sliding gate. The bottom level of the channel is 0.33 m higher than the reservoir bottom level,

with a vertical step at the channel inlet (see figure 20). In the experiment, the gate is pulled up very quickly and the closure failure is assumed as instantaneous. The channel is equipped with a set of measurement devices; their location is shown in table 2. The walls of the channel are made of glass and the bed is made of steel; the n Manning friction coefficient has been assumed equal to 0.0095 s/m$^{1/3}$ and the wall friction effect has been neglected.

The initial water level in the upstream reservoir is 0.2 m above the channel bed level and the corresponding water depth is equal to 0.53 m. The downstream channel is dry in a first set of measurements and wet with a water depth 0.01 m in a second one.

When the gate is opened, the water flows rapidly into the channel and reaches the bend after approximately 3 s. There, the water reflects against the wall, a bore forms and begins to travel in the upstream direction, back to the reservoir. For the water flowing downstream after the bend, multiple reflections on the walls can be observed.

The number of elements of the mesh used by the MAST algorithm is 10919 and the mean value of the element area is 8.4 E-04 sm; time step Δt is 0.01 s.

In figures 21,a and 21,b the contour plot of the water depths is shown for the simulation times 3 s and 6 s for the first set of experiments (dry bed). Multiple reflections are evident in the channel downstream the bend. Observe also that no numerical oscillations of the water level occur in the transition reservoir-channel, even though the shallow water hypothesis does not hold at the vertical step.

In figures 22,a-22,f and 23,a-23,f the comparison between measured and computed water depths at measure points is shown for the two sets of experiments. The maximum CFL numbers are respectively 1.71 and 2.26.

In the figures, results computed in both dry and wet bed conditions by *Zhou et al*. (2002) using the SGMS algorithm and, in dry bed conditions only, by *Gottardi and Venutelli* (2004) are shown. *Zhou et al*. (2002) used a regular quadrilateral mesh with Δx = 0.05017 m and Δy = 0.495 m. *Gottardi and Venutelli* (2004) discretized the domain using a regular quadrilateral mesh with side $\Delta x = \Delta y$ = 0.01 m.

At gauge G1 inside the reservoir, in both wet and dry conditions, all the numerical schemes produce very similar results and for simplicity only the ones provided by the MAST scheme are reported against the experimental data.

Observe the time delay of the shock of the reflected wave in both MAST and Gottardi and Venutelli' s results with respect to the measured data at gauges G2, G3 and G4. For MAST scheme this delay at gauge G4 is approximately 1.8 s in dry bed conditions and 1.55 s for wet bed conditions. The delay reduces progressively travelling from gauge G4 to gauge G2. In dry bed conditions this delay is about 1.2 s, while in wet bed conditions the effect of the reflected wave in the numerical results arrives early (about 0.2 s) with respect to the measured data. Results by *Zhou et al*. (2002) are in good agreement with measured data at gauge G3, while the effect of the reflected wave shows a little delay (about 1 s) at gauge G4 and arrives early at gauge G2 (about 1.7 s and 1.45 s in dry and wet bed conditions). At gauge G2 both MAST and SGMS overestimate water depth before the arrival of the reflected wave and results are very similar. On the opposite, *Gottardi and Venutelli* (2004) underestimate water depth approximately up to 10 s.

At gauges G3 and G4, the three schemes provide similar result before and after the arrival of the reflected wave. After the arrival of the reflected wave, the three numerical schemes produce similar results also at the gauge G2.

Differently from *Zhou et al*. (2002), experimental data at point G5 are properly simulated by the MAST scheme in both wet and dry conditions. At gauge G6 all the numerical results are similar to the measured data.

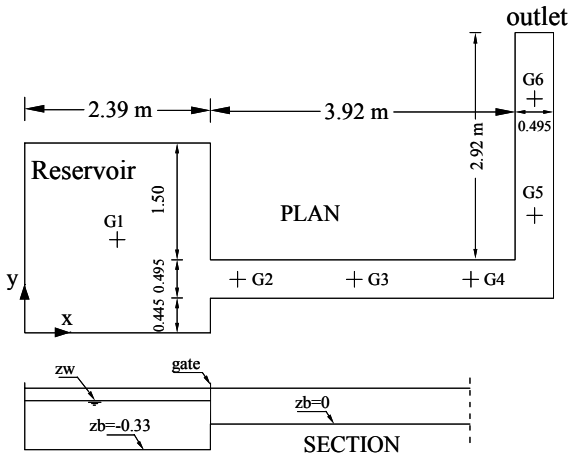

Figure 20. Plane view and profile of the entrance of the L-shaped channel

Point	*x* [m]	*y* [m]
G1	1.19	1.2
G2	2.74	0.69
G3	4.24	0.69
G4	5.74	0.69
G5	6.56	1.51
G6	6.56	3.01

Table 2. Spatial coordinates of the experimental points for test 3.2

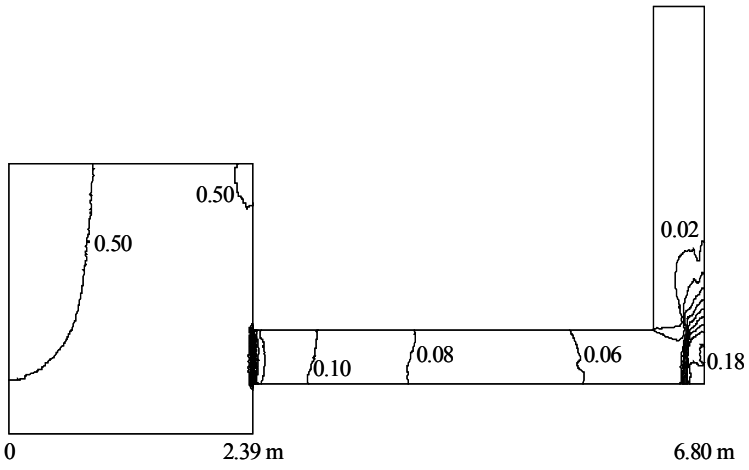

Figure 21,a. Contour plot of the water depths after 3 s (dry bed)

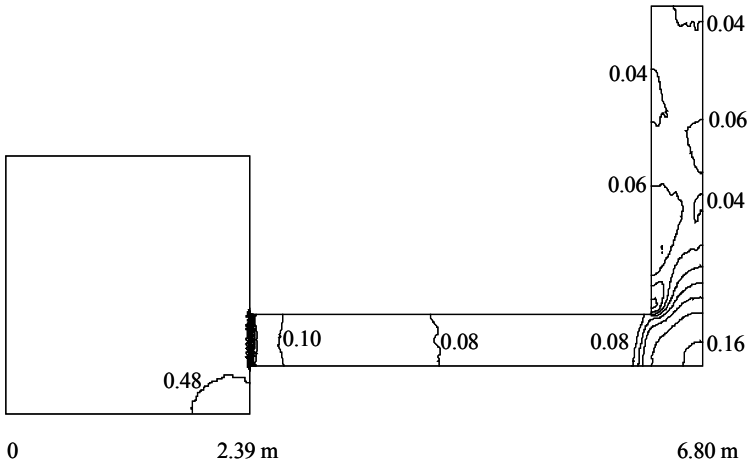

Figure 21,b. Contour plot of the water depths after 6 s (dry bed)

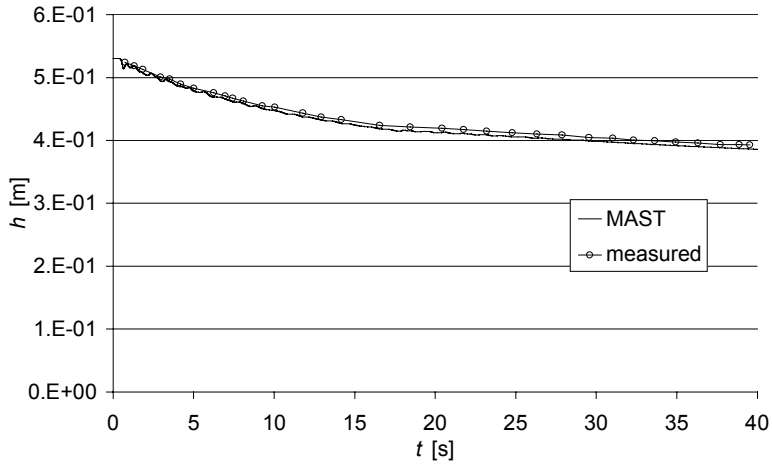

Figure 22,a. Water depths at the experimental points G1 (dry bed)

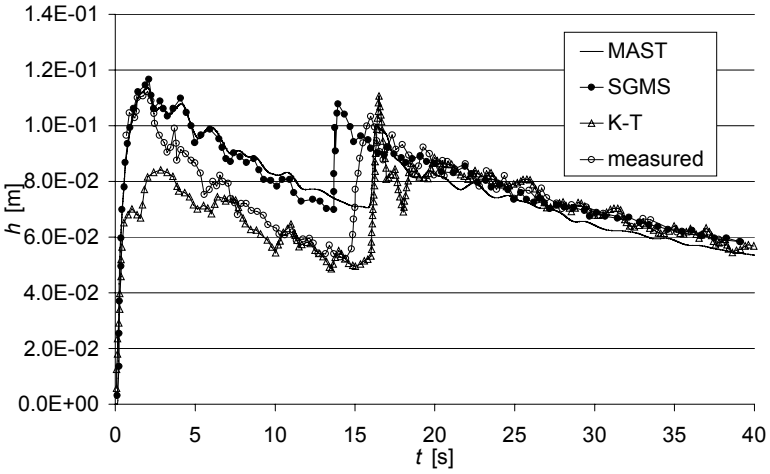

Figure 22,b. Water depths at the experimental points G2 (dry bed)

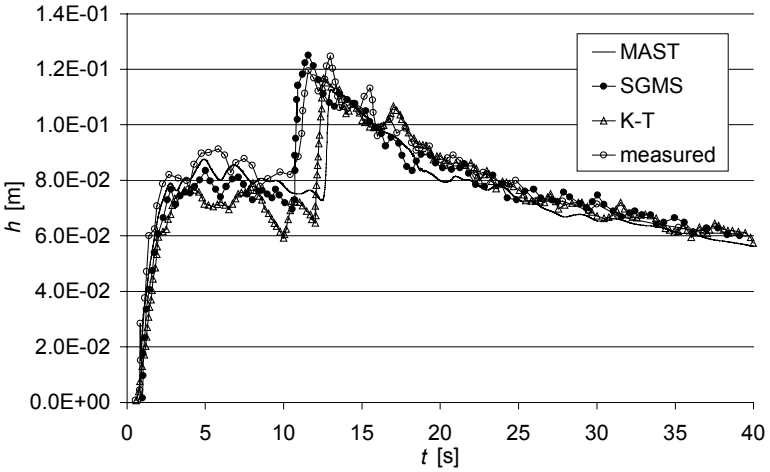

Figure 22,c. Water depths at the experimental points G3 (dry bed)

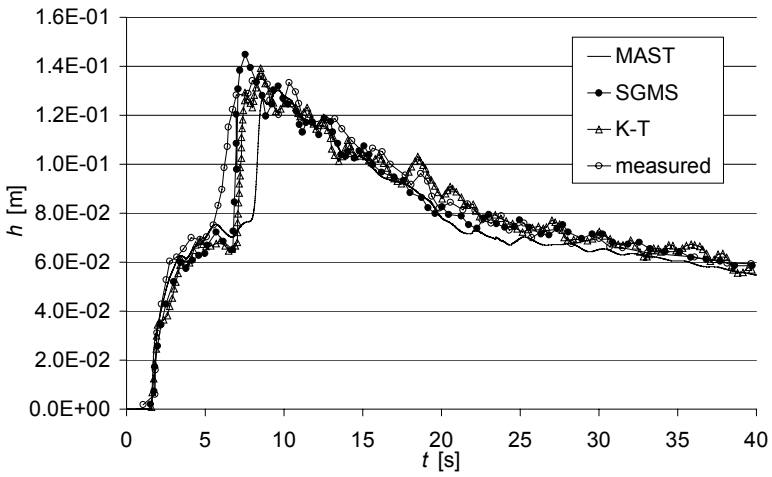

Figure 22,d. Water depths at the experimental points G4 (dry bed)

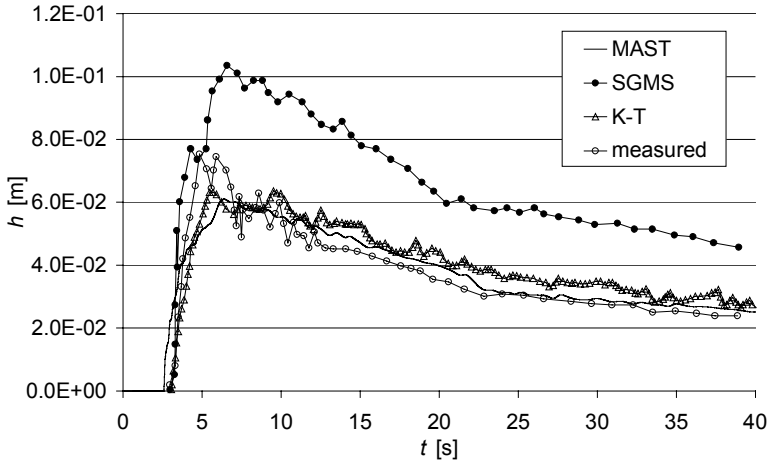

Figure 22,e. Water depths at the experimental points G5 (dry bed)

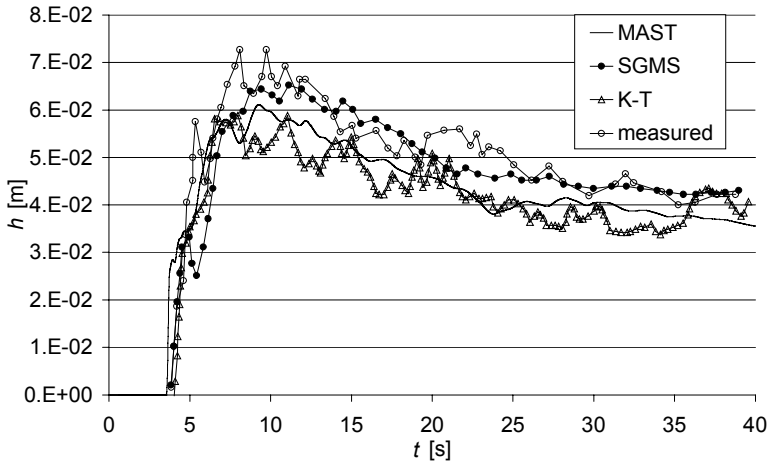

Figure 22,f. Water depths at the experimental points G6 (dry bed)

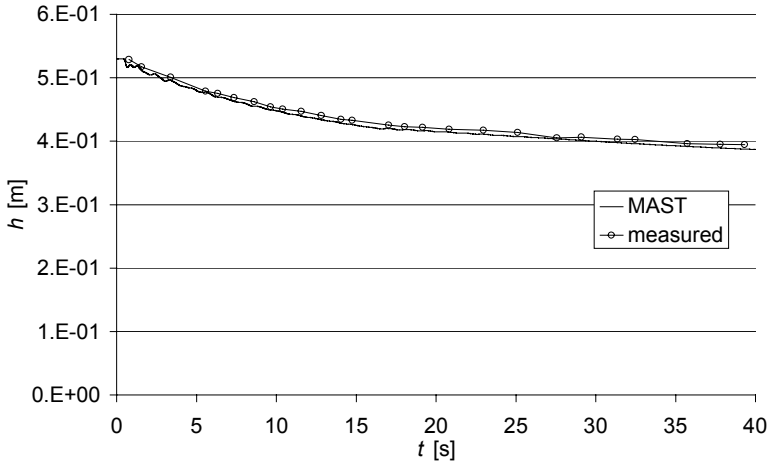

Figure 23,a. Water depths at the experimental points G1 (wet bed)

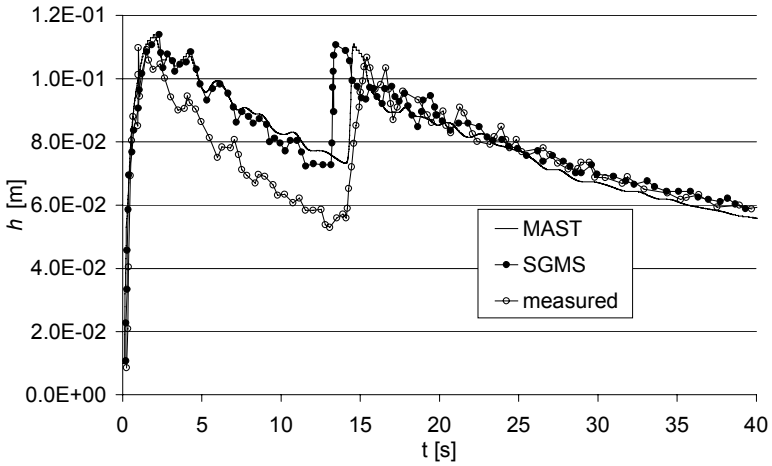

Figure 23,b. Water depths at the experimental points G2 (wet bed)

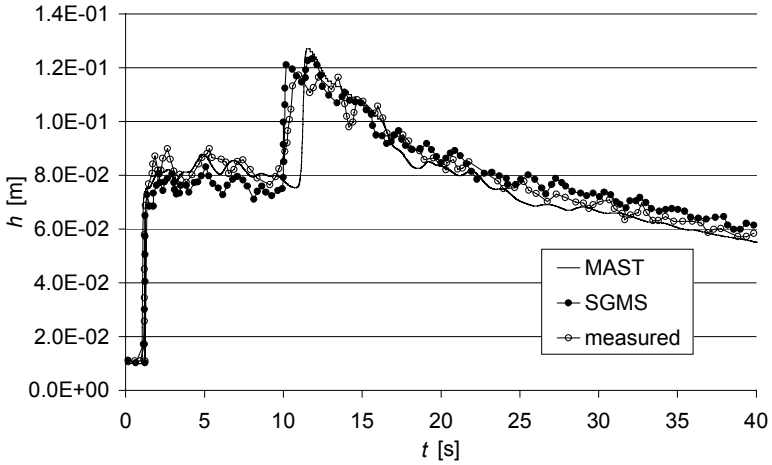

Figure 23,c. Water depths at the experimental points G3 (wet bed)

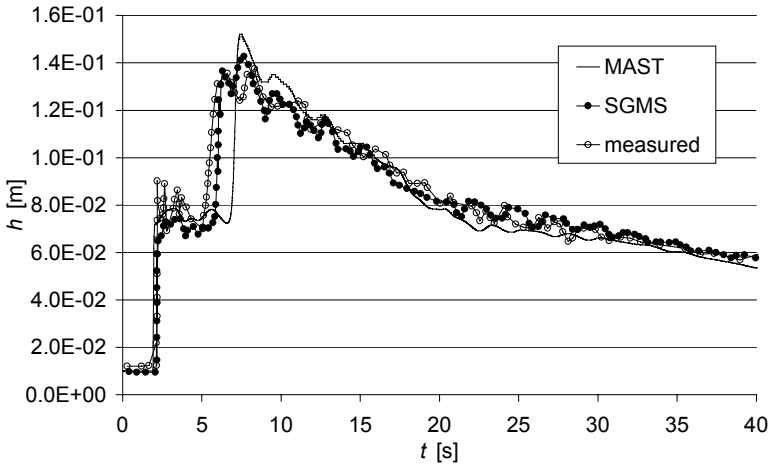

Figure 23,d. Water depths at the experimental points G4 (wet bed)

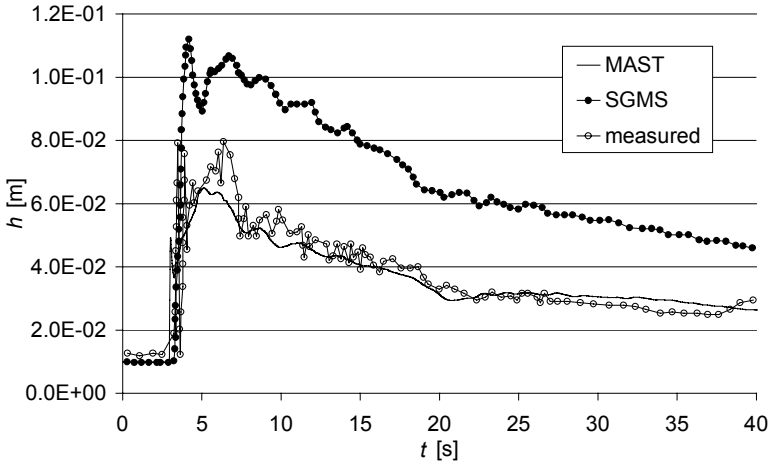

Figure 23,e. Water depths at the experimental points G5 (wet bed)

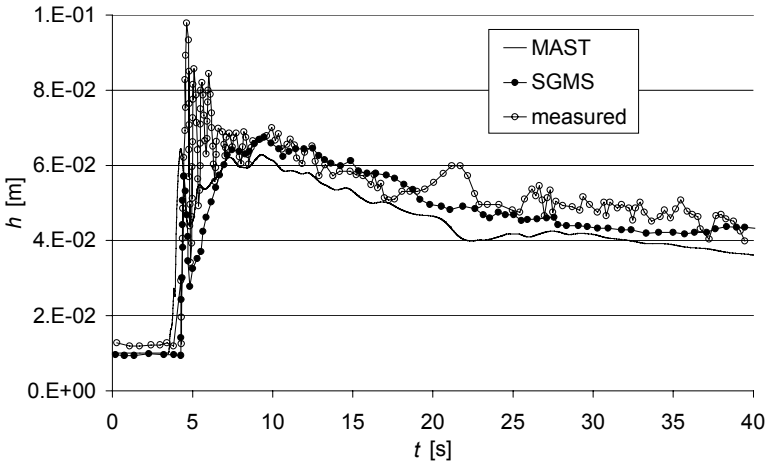

Figure 23,f. Water depths at the experimental points G6 (wet bed)

6.3 CPU times investigation

CPU mean time values for the computation of the convective and diffusive steps, as well as for the computation of the approximated potential scalar function for each element are reported in table 3 for all the numerical tests. For the computations, a Pentium 4, CPU 2GHz processor has been used. In table 3 CPU times are in seconds.

Observe that in this case the computation of the convective prediction and correction steps is the most demanding one; the CPU time required for the computation of the diffusive step is approximately 1 magnitude order less that the convective steps time. The CPU time for the computation of the approximated scalar function is about half the convective steps time.

The growth of the computational time per element and per time step versus the number of elements has been investigated, in order to assess the relative cost of the diffusive correction step and of the approximated potentials calculation. These steps require the solution of a large linear system and are the non 'explicit' component of the algorithm. The initial and boundary conditions of the experimental test of Fraccarollo and Toro [30] with zero bottom slope have been used. A quite coarse irregular mesh has been generated and used as the basis for further refinements. At each refinement level $i + 1$, each triangle of the previous i^{th} mesh is subdivided in four equal triangles (see figure 24). In figure 25 the 1^{st}-level mesh with 372 elements is shown. Δt is 0.04 s for the 1^{st}-level mesh; for the next level meshes time step size has been halved, in order to limit the growth of the maximum CFL number.

For each mesh density, the α_n normalized parameter, weighting the penalty function in the functional F' of Eq. (93) and defined as

$$\alpha_n = \frac{\alpha}{\sum_j 2\left|FL_{e,j}^k\right|} \tag{114},$$

has been changed from 0 to the maximum value affecting for no more than 1 per cent the computed water depths. This maximum value has been empirically found to be 0.01.

Results are shown in figures 26,a-26,c. Observe that, at least up to the maximum number of elements tested in the experiments, the CPU components relative to the convective and diffusive steps can be considered almost independent from the elements number. The small decrement of the average CPU time of the convective component can be related to the increasing CFL numbers obtained by partitioning and

to the best aptitude of the algorithm to work with CFL numbers greater than one [27].

The choice of the α_n coefficient is very important for the computational cost. The structural difference between the linear system associated to the diffusive problem and the system associated to the approximated potential estimation is that the extra diagonal term is given in the first one by the local time derivative and in the second one by the penalty term. This implies that the condition number of the second one increases along with the size of the α_n coefficient. On the other hand, an excessive value can ruin the potential quality.

Test	potential	convective steps	diffusive step
3.1 - 0 slope	4.982E-05	8.800E-05	1.428E-05
3.1 - 0.07 slope	5.045E-05	9.060E-05	1.472E-05
3.2 - dry bed	5.741E-05	8.126E-05	1.499E-05
3.2 - wet bed	5.591E-05	8.042E-05	1.512E-05

Table 3. CPU time values

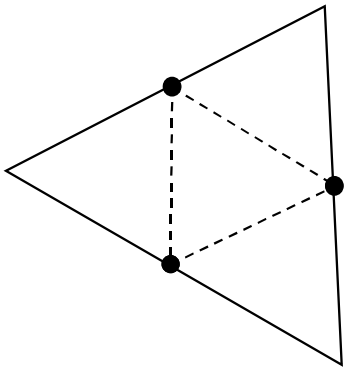

Figure 24. Mesh refinement

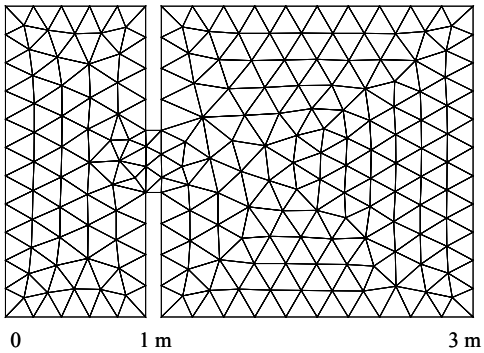

Figure 25. The 1st-level mesh for the CPU time investigation

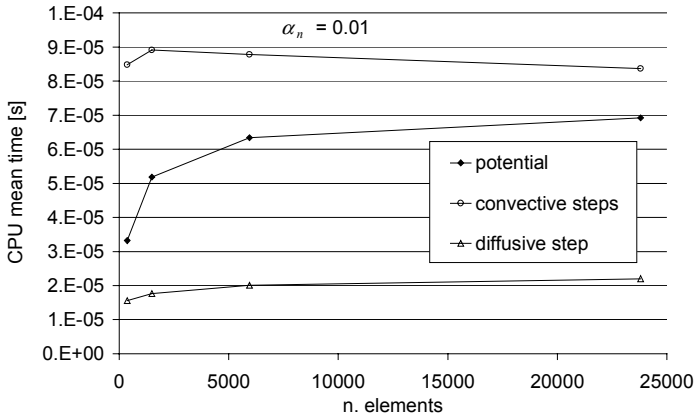

Figure 26,a. Effect of the element number on the CPU time; α_n = 1.E-02

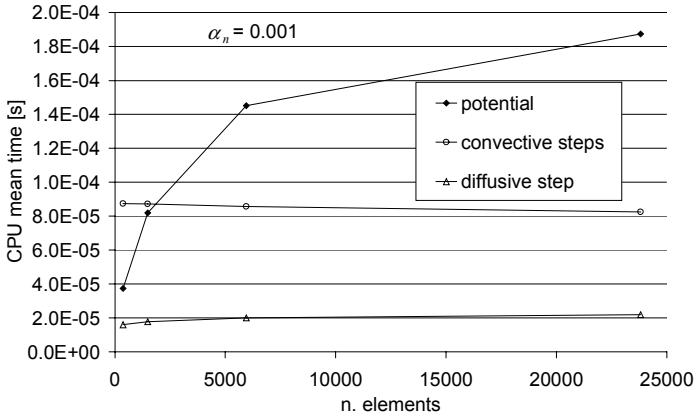

Figure 26,b. Effect of the element number on the CPU time; α_n = 1.E-03

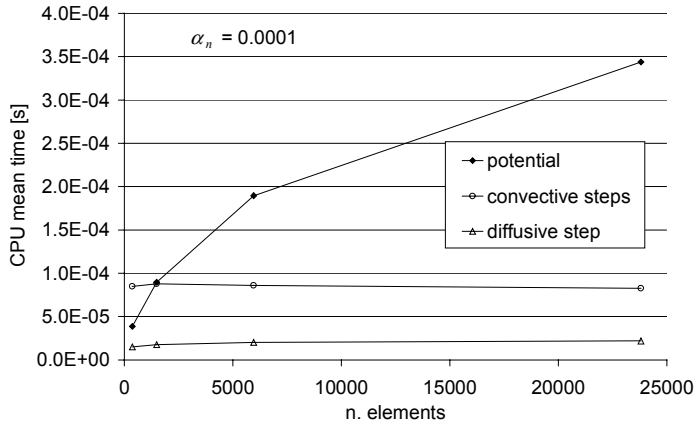

Figure 26,c. Effect of the element number on the CPU time; α_n = 1.E-04

The sensitivity of the numerical solution with respect to the CFL number has been investigated by changing the size of the time step. The test carried out by Fraccarollo and Toro [30] with zero bottom slope has been chosen again for this study. The computational mesh shown in figure 15 has been used. In figure 27 the contour plot of the water level after 0.5 s are shown; these numerical simulations have been carried out using Δt = 0.005 s and Δt = 0.025 s. The maximum CFL numbers are respectively 1.11 and 5.35. These results have to be compared also with similar ones shown in figure 16 (right) and computed with a maximum CFL number about 3.62. The test shows the result sensitivity with respect to the CFL number to be very small.

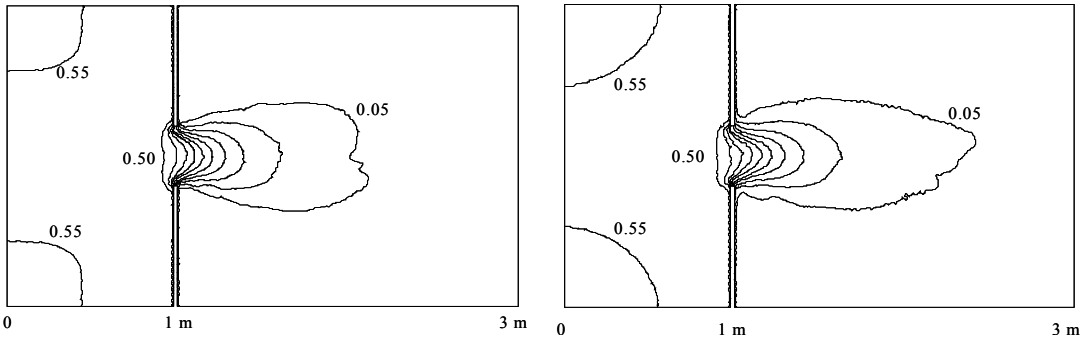

Figure 27. Contour plots of the water level after 0.5 s – test 3.1, CFL 1.11(left) and 5.35 (right)

CONCLUSIONS

A new algorithm for the solution of the complete 1D and 2D Saint-Venant equations has been presented. The original PDEs system is split in three different ones. Each of the first two resulting systems is equivalent to a scalar convective problem and is solved using a MAST technique. This technique solves the convective problem with a sequential solution of small ODEs system, with order equal to the number of the time-dependent variables per computational cell. A self-adjusting step size algorithm is used for the solution of each ODEs system from the known time level to the next one. Results show no restriction on the size of the original time step and of the source terms. The third PDEs system is similar to the original one, but the volume and momentum fluxes, along with the computed dependent variables changes, are quite small. This allows a linearization of the problem, that is solved using a finite difference fully implicit discretization, leading to the solution of a very well conditioned linear system with order equal to the cell number.

According to the tests carried out in this paper, the algorithm can be classified as 'explicit', because the main computational effort is required by the solution of the convective problems, and is almost proportional to the number of elements. The algorithm components that have a computational cost more than linearly increasing with the number of elements are 1) the solution of the linear system in the diffusive correction step and 2) the ordering of the elements required for the application of the MAST technique. In the proposed 1D tests, the average ratio between the CPU time required by these parts of the algorithm an the total CPU time is about 0.05.

In the 2D problem, the computational time of the third problem has been always the smaller one in all the performed simulations. The larger CPU time is actually dedicated, for larger elements number, to the computation of the approximated potential; even with a negligible weight of the penalty term the increment of the CPU specific time, per element and per time step, is much less than linear. A normalized value of this term equal to 0.01 seems to provide almost constant specific CPU time without a meaningful loss of accuracy.

Model results are in good agreement with analytical solutions and lab experiments. Even if a first order spatial approximation has been used for the dependent variables, the achieved accuracy is similar to the accuracy obtained by higher order approximation schemes, using the same mesh density and time step size. MAST performance is even better if the time step and the corresponding maximum CFL number is increased beyond the limit of one required by the other explicit schemes, with a corresponding reduction of the CPU time. The optimal CFL number, for the MAST algorithm, is usually between one and five, and a strong stability of the results can be observed changing the time step size. Good results are also obtained in terms of flow rates, especially where strongly variable source terms occur, like in the test 4.2 (steady flow over a bump).

REFERENCES

Abbott M. B. and Minns A. "*Computational Hydraulics*". Ashgate ed., (1992) UK.
Alcrudo F. and Garcìa – Navarro P. "*A high – resolution Godunov – tipe scheme in finite volume for the 2D shallow water equations* ". Int. J. Num. Meth. Fluids, vol 14 (1993), pp. 489-505.
Anastasiou K. and Chan C. T. "*Solution of the 2D shallow water equations using the finite volume method on unstructured triangular mesh*". Int. J. Num. Meth. Eng., vol. 24(11) (1997), pp. 1225-1245.
Argus Holdings Ltd.
http://www.argusint.com/pub/OnLineDocs/ArgusONE_UsersGuide.pdf
Aricò C. and Tucciarelli T. "*MAST solution of advection problems in irrotational flow fields*". In press for Adv. Wat. Res. (2006)
Aricò C., Nasello C. and Tucciarelli T. "*A marching in space and time solver for the complete 2D shallow water equations. Application to real test cases*". Proc. Int. Conf. River Flow 2006, 6-8 Sept. (2006) Lisbon.
Bascià A. and Tucciarelli T. "*An explicit unconditionally stable numerical solution of the advection problem in irrotational flow fields*" . Wat. Res. Res., Vol. 40, n. 6 (2004).
Berger M. J., Aftomis M. J. and Melton J. E. "*Adaption and surface modelling for Cartesian mesh methods*". AIAA (1995)-1725-CP, pp. 881-972.
Bermudez A. and Vazquez – Cendon M. E. "*Upwind methods for hyperbolic conservation laws with source terms*". Comput. Fluids vol 23(8) (1994), p. 1049.
Causon D. M., Ingram D. M., Mingham C. G., Yang G., and Pearson R.V. "*Calculation of shallow water flows using a Cartesian cut cell approach*". Adv. Wat. Res., vol. 23 (2000), pp. 545-562.
Chiang Y., Van Leer B. and Powell K. G. "*Simulation of unsteady inviscid flow on an adaptively refined Cartesian grid*". AIAA (1992) 92-0443.
De Saint Venant B. "*Thèorie de mouvement non-permanant des eaux avec application aux crues des rivières et à l'introduction des marèes dans leur lit*". Acad. Sci. Paris C. R., 73 (1871), pp. 148-154, 237-240.
De Zeeuw D. and Powell K. G. "*An adaptively refined Cartesian mesh solver for the Euler equations*". J. Comp. Phys., vol. 104(1) (1993), pp. 56-124.
Fraccarollo L. and Toro E. F. "*Experimental and numerical assessment of the shallow water model for two-dimensional dam-break type*". J. Hydr. Res., vol. 33(6) (1995), pp. 843-864.
Gottardi G. and Venutelli M. "*Central scheme for dam-break flow simulation*". Adv. Wat. Res., vol. 27 (2004), pp. 259-268.
Gordon W. J. and Hall C. A. "*Construction of curvilinear co-ordinate systems and applications to mesh generation*". Int. J. Num. Meth. Eng., vol. 7 (1973), pp. 461-538.
Hauser J., Paap H. G., Eppel D. and Mueller A. "*Solution of the shallow water equations for complex flow domains via boundary-fitted coordinates*". Int. J. Num. Meth. Fluids, vol. 5 (1985), pp.727-768.
Hubbard M. E. "*Multidimensional slope limiters for MUSCL-type finite volume schemes on unstructured grids*". J. Comp. Phys., vol. 155(1) (1999), pp. 54-74.
Hubbard M. E. and Garcìa – Navarro P. "*Flux difference splitting and the balancing of the source terms and flux gradients*". J. Comp. Phys., vol. 165(89) (2000), pp. 89-125.
Jiwen W. and Ruxun L. "*The composite finite volume method on unstructured meshes for the two-dimensional shallow water equations*". Int. J. Num. Meth. Eng., vol. 37(8) (2001), pp. 933-949.
Kurganov A. and Tadmor E. "*New high-resolution central schemes for non-linear conservation laws and convection-diffusion equations*". J. Comp. Phys. vol. 160 (2000), pp. 241-282.
LeVeque R. J. "*Balancing flux terms and flux gradients in high-resolution Godunov methods: the quasi steady wave propagation algorithm*". J. Comp. Phys., vol. 146(1), (1998), p. 346.
Meselhe E. A., Sotiropoulos F. and Holly F. H. "*Numerical simulation of transcritical flow in open channels*".ASCE J. Hydr. Eng., vol. 123(9) (1997), pp. 774-783.
Mingham C. G. and Causon D. M. "*High – resolution finite – volume method for shallow water flows*". ASCE J. Hydr. Eng., vol. 124(6) (1998), pp. 605-614.
Nag Library Manual (2005) http://www.nag.co.uk/numeric/fl/manual/html/mark21.html
Noto V. and Tucciarelli T. "*DORA Algorithm for network flow models with improved stability and convergence properties*". ASCE J. Hydr. Eng., vol. 127 (5) (2001), pp. 380-391.
Putti, M., Yeh W. W. G. and Mulder W. A. "*A triangular finite volume approach with high resolution upwind terms for the solution of groundwater transport equations*". Wat. Res. Res., vol. 26(12) (1990), pp. 2865-2880.
Ritter A. (1992). „*Die Fortplanzung der wasserwellen*". Zeitschrift des Vereines Deutscher Ing., vol. 36(33), pp. 947-954.
Shu C. W. and Osher S. "*Efficient implementation of essentially non-oscillatory shock capturing schemes*". J. Comp. Phys., vol. 77 (1988), pp. 439-471.
Soares Frazão S., Sillen X. and Zech Y. Dam-break flow through sharp bends physical model and 2D Boltzmann model validation. Wallingford CADAM meeting, 2-3 March, 1998.
http://www.hrwallingford.co.uk/projects/CADAM/CADAM/index.html .

Stoker J. *Water Waves*. Intersciences (1957).

Thompson J. F., Thames F. C. and Mastin C. W. "*Automatic generation of body-fitted curvilinear coordinate system for field containing any number of arbitrary two-dimensional bodies*". J. Comp. Phys., vol. 15 (1974), pp. 299-618.

Toro E. F. "*A weighted averaged flux method for hyperbolic conservation laws*". *Proc. R. Soc. Lond*. A423 (1989), pp. 401-418.

Toro E. F. "*Riemann solvers and numerical methods for fluid dynamics. A practical introduction*". Springer 1st edition, (1997).

Tucciarelli T. "*A new algorithm for a robust solution of the fully dynamic Saint – Venant equations*". J. Hydr. Res., vol. 41(3) (2003), pp. 239-246.

U.S. Army Corp of Engineers (USACE). "*Flood routing from suddenly breached dams*". Miscellaneous Papers vol. 2 (374) Rep. 1 (1960), U.S. Army Engineer Waterways Experiment Station, Vicksburg, Miss.

Vukovic S. and Sopta L. "*ENO and WENO schemes with exact conservation property for one – dimensional shallow water equations*". J. Comp. Phys., vol. 179 (2002), pp. 593-621.

Yang J. and Hsu C. A. "*Computation of free surface flows*". IAHR J. Hydr. Res., vol. 31(3) (1993), pp. 403-416.

Yoon T. H. and Kang S. K. "*Finite Volume Model for Two-Dimensional Shallow Water Flows on Unstructured Grids*". ASCE J. Hydr. Eng., vol. 130(7) (2004), pp. 678-688.

Zhang S. Q., Ghidaoui M. S., Gray W. G. and Li N. Z. "*A kinetic flux vector splitting scheme for shallow water flows*". Adv. Wat. Res. Vol. 26 (2003), pp. 635-647.

Zhou J. G., Causon D. M., Mingham C. D. and Ingram D. M. "*The surface gradient method for the treatment of the source term in the shallow water equations*". J. Comp. Phys., vol. 168 (2001), pp. 1-25.

Zhou J. G., Causon D. M., Ingram D. M. and Mingham C. D. "*Numerical solutions of the shallow water equations with discontinuous bed topograph*y". Int. J. Num. Meth. Fluids vol. 38 (2002), pp. 769-788.

Zhou J. G., Causon D. M., Mingham C. G. and Ingram D. M. "*Numerical prediction of dam-break flows in general geometries with complex bed topography*". ASCE J. Hydr. Eng., vol. 130(4) (2004), pp. 332-340.

Zic N. C., Vukovic S. and Sopta L. "*Balanced finite volume WENO and central WENO schemes for the shallow water and the open – channel flows equations*". J. Comp. Phys. vol. 200, (2004), pp. 512-548.

Zoppou C. and Roberts S. Discussion of: "*Space-Time conservation methods applied to Saint Venant equations*" (May 1998, by Molls T. and Molls F.), ASCE J. Hydr. Eng. vol. 125(8) (1999), pp. 891-892.

Zoppou C. and Roberts S. "*Explicit schemes for dam – break simulations*". ASCE J. Hydr. Eng., vol. 129 (1) (2003), pp. 11-34.