A well-balanced spectral volume scheme with the wetting–drying property for the shallow-water equations
Luca Cozzolino, Renata Della Morte, Giuseppe Del Giudice, Anna Palumbo and Domenico Pianese

ABSTRACT
The shallow-water equations are widely used to model surface water bodies, such as lakes, rivers and the swash zone in coastal flows. Physically congruent solutions are characterized by non-negative water depth, and many numerical methods may fail to preserve this property at the discrete level when moving wet–dry transitions are present in the physical domain. In this paper, we present a spectral-volume method for the approximate solution of the one-dimensional shallow-water equations, which is third-order accurate in wet regions, far from discontinuities, and which is well balanced for water at rest states: the stability of the solution is ensured if reconstruction and limitation of variables preserves non-negativity of the depth and a suitable constraint for the time step length is satisfied. A number of numerical experiments are reported, showing the promising capabilities of the model to solve problems with non-trivial topographies and friction.

Key words | friction, high-order accuracy, shallow-water equations, spectral volume method, well-balancing property, wetting-drying property

INTRODUCTION
The increase of safety demand against adverse flooding phenomena, and the request of better understanding of propagation of constituents in surface water bodies, have prompted the development of many mathematical models for the description of shallow flows in lakes, rivers and the swash zone of coastal areas. When the water depth is negligible with respect to the horizontal dimensions of surface water bodies, long waves can be efficiently described by means of the so-called shallow-water equations. Since the analytical solution of such mathematical models is available only in very schematic cases, numerical methods are required in general cases.

In one-dimensional gradually varied flows, the shallow-water equations can be obtained from the De Saint Venant equations (Cunge et al. 1980) by assuming a rectangular cross section with unitary width: despite these simplification, the one-dimensional shallow-water equations retain all the mathematical characteristics of the De Saint Venant equations and are universally considered a model for developing and testing numerical methods useful for the solution of the complete system of governing equations. Moreover, the one-dimensional shallow-water equations are interesting per se, in that their solution is the fundamental ingredient for the numerical approximation of two-dimensional shallow-water equations, which are used to simulate the flow characteristics in two-dimensional shallow-water bodies (Toro 2001). In one space dimension, the shallow-water equations assume the form

\[
\begin{align*}
\frac{\partial h}{\partial t} + \frac{\partial h U}{\partial x} &= 0 \\
\frac{\partial h U}{\partial t} + \frac{\partial}{\partial x} \left( \frac{1}{2} g h^2 + h U^2 \right) &= -g h \frac{dz_b}{dx} - g h S_f
\end{align*}
\]

where the following definitions hold: \(z_b(x)\) = bed elevation, \(h(x, t)\) = water depth, \(U(x, t)\) = vertically averaged water
velocity, \( g \) = gravity acceleration constant, \( S_f (h, U, x) \) = friction slope, \( x \) = horizontal coordinate, \( t \) = time. The first of Equation (1) represents the mass conservation, while the second represents the momentum balance. Since these hyperbolic equations are nonlinear, they admit the formation of discontinuities in the flow field (propagating bores, hydraulic jumps) even in the case of smooth initial conditions: such discontinuities are also encountered in the case of propagation of wetting–drying moving frontiers. Different numerical schemes, based on the concept of finite differences (FD), finite volumes (FV) or finite elements (FE), when written in conservative form, with numerical fluxes evaluated using an exact or approximate Riemann solver (Toro 2001), are able to cope with discontinuities in the flow field (Mynett 1999).

We observe that the momentum equation is characterized by the presence of source terms, due to the bed slope and friction, on the right-hand side: this term balances the advective fluxes when a steady-state condition is attained. The numerical solution of the shallow-water equations with source terms has to take into account this fact, in order to avoid non-physical solutions, difficulties in attaining steady state conditions, spurious oscillations of the numerical approximation and instability of the algorithm. When water is at rest, the stationary state reduces to

\[
U = 0; \quad \zeta = \text{const.} \tag{2}
\]

where \( \zeta = h + z_b \) is the water surface elevation. Many classic FV schemes may fail in guaranteeing the mutual cancelation of fluxes and source terms when the water is at rest is the physically congruent solution: the numerical methods which preserve the water-at-rest condition are said to satisfy the so-called \( C \)-property (Bermúdez & Vázquez 1994), and a number of different source term treatments which ensure this property have been proposed in the literature (see Caselles et al. 2009 for a brief review).

Not properly designed schemes for the solution of the shallow-water equations produce negative depths near wet–dry transitions, which can reduce the accuracy of the method or introduce stability issues: physically admissible solutions of Equation (1) are characterized by \( h \geq 0 \), and it is desirable that the numerical schemes are able to mimic this non-negativity property at the discrete level. Also, in the case that possible small negative depths are set to zero at every time step, the mass error accumulates: in the worst cases, the appearance of negative depths plainly leads to the instability of the scheme. Numerous second-order accurate FV schemes, able to preserve water-at-rest states and non-negativity of the solution (Kurganov & Levy 2002; Gallardo et al. 2007; Berthon & Marche 2008), have been presented.

Despite the fact that high-order FV schemes have been available in the numerical modelling community for a long time (Colella & Woodward 1984), first- and second-order schemes have been generally preferred when the approximate solution of the shallow-water equations has been considered (Audusse & Bristeau 2005): in fact, low-order schemes are robust and simple estimates can be given for the stability requirements (Bouchut 2004) and they can be easily implemented also in the case of unstructured grids. Of course, first- and second-order FV schemes exhibit a major shortcoming, namely high diffusive and dispersive errors: this could not seem an issue in the case of fast transients (dam break or run-up evaluation), but in the case of long transients, or in the case of small-amplitude wave propagation, the quality of the solution can be adversely affected (Shu 2005). This justifies the recent effort made by researchers in order to implement genuinely high-order FV schemes for the approximate solution of the shallow-water equations (Caleffi et al. 2006; Castro et al. 2006; Noelle et al. 2006; Xing & Shu 2006a, b; Akoh et al. 2008): generally, these schemes do not exhibit the wetting–drying property. In a very recent work (Xing & Shu 2011), a high-order finite-volume scheme is presented, able to manage dry or nearly dry regions, but which seems unable to cope with friction source terms.

In this paper we describe and demonstrate a novel third-order accurate well-balanced numerical model for the approximate solution of the one-dimensional shallow-water equations with bed-slope and friction source terms, named SV3RK3, which is able to cope with moving wet–dry frontiers, and also when friction and complex topography are present. In the following sections, the treatment of the bed-slope and friction source terms is presented in the context of the spectral volume (SV) method. The modifications needed to satisfy water-at-rest conditions in the
THE 1D SPECTRAL VOLUME METHOD FOR THE SHALLOW-WATER EQUATIONS

The SV method is a high-order numerical method, recently developed for the solution of conservation laws on unstructured grids (Wang 2002; Wang & Liu 2002, 2004; Wang et al. 2004; Zhang & Shu 2005; Liu et al. 2006; Sun et al. 2006; Van den Abeele & Lacor 2007; Van den Abeele et al. 2007a, b), which is strongly related to the spectral difference (SD) method, to the discontinuous Galerkin (DG) method and to the FV method: the solution is approximated by means of piecewise continuous polynomials in each element, as in the DG and SD methods, while the degrees of freedom of the reconstruction are represented by the averaged values of the conserved variables in the FV which constitute a SV. The SV method has been applied to the solution of the shallow-water equations, both discarding (Choi et al. 2004) or introducing (Cozzolino & Pianese 2006) the source terms. In the present section, we describe a third-order SV method for the solution of the shallow-water equations, which preserves the C-property and which is able to cope with wet–dry moving frontiers.

Basics of the one-dimensional SV method

Consider the shallow-water equations system

\[ \frac{\partial \mathbf{u}}{\partial t} + \frac{\partial \mathbf{f}(\mathbf{u})}{\partial x} = \mathbf{s}_0(\mathbf{u}, x) + \mathbf{s}_f(\mathbf{u}, x) \]

(3)

where \( \mathbf{u} \) is the column vector of the conserved variables, whose generic component is \( u \), \( \mathbf{f}(\mathbf{u}) \) is the column vector of the fluxes, while \( s_0(\mathbf{u}, x) \) and \( s_f(\mathbf{u}, x) \) are the column vectors of the bed-slope and friction source terms, respectively. The following definitions hold:

\[ \mathbf{u} = \begin{bmatrix} h & hU \end{bmatrix}^T, \quad \mathbf{f}(\mathbf{u}) = \begin{bmatrix} hU & \frac{1}{2}gh^2 + hU^2 \end{bmatrix}^T, \]

\[ \mathbf{s}_0(\mathbf{u}, x) = \begin{bmatrix} 0 & -gh \frac{dz_b}{dx} \end{bmatrix}^T, \]

\[ \mathbf{s}_f(\mathbf{u}, x) = \begin{bmatrix} 0 & -ghS_t \end{bmatrix}^T \]

(4)

and \( T \) is the symbol of ‘matrix transpose’. The one-dimensional physical domain \( \Omega = [a, b] \) is partitioned into \( NS \) non-overlapping cells named ‘spectral volumes’ or ‘spectral cells’, indexed by \( i \):

\[ \Omega = \bigcup_{i=1}^{NS} S_i, \quad S_i = [x_{i-1/2}, x_{i+1/2}], \quad i = 1, 2, \ldots, NS \]

(5)

with \( x_{1/2} = a \) and \( x_{N+1/2} = b \). In order to attain the third order of accuracy for the space discretization, the spectral cells are in turn partitioned by means of three non-overlapping FV, indexed by \((i, j)\):

\[ S_i = \bigcup_{j=1}^{3} C_{ij}, \quad C_{ij} = [x_{ij-1/2}, x_{ij+1/2}], \]

(6)

where \( x_{i,1/2} = x_{i-1/2} \) and \( x_{i,3+1/2} = x_{i+1/2} \) (see Figure 1).
Let $\mathbf{u}_i$ be the cell-averaged value of the conserved variables vector $\mathbf{u}$ in $S_i$ and $\mathbf{u}_{i,j}$ the FV-averaged value of $\mathbf{u}$ in $C_{i,j}$. The following obvious relation between $\mathbf{u}_i$ and $\mathbf{u}_{i,j}$ holds:

$$
\mathbf{u}_i = \frac{1}{\Delta x_i} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{u} \, dx = \frac{1}{\Delta x_i} \sum_{j=1}^{3} \int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{u} \, dx = \frac{1}{\Delta x_i} \sum_{j=1}^{3} \mathbf{u}_{i,j} \Delta x_{i,j}
$$

where $\Delta x_i$ is the length of the SV $S_i$, while $\Delta x_{i,j}$ is the length of the FV $C_{i,j}$. Congruently with the finite-volume representation of conserved variables, defined by Equation (7), the bed elevation is represented in each FV $C_{i,j}$ by its finite-volume-averaged value $z_{i,j}$.

Applying the method of lines, the balance laws (3) are spatially integrated in each FV and the Green theorem yields the following system of ordinary differential equations:

$$
\frac{d\mathbf{u}_{i,j}}{dt} = -\frac{1}{\Delta x_{i,j}} (f_{i,j+1/2} - f_{i,j-1/2}) + \mathbf{s}_{b_{i,j}} + \mathbf{s}_{f_{i,j}}
$$

$$
i = 1, 2, \ldots, \text{NS}, \quad j = 1, 2, 3
$$

where $f_{i,j+1/2}$ is the vector of the numerical fluxes at the interface between the FV $C_{i,j}$ and $C_{i,j+1}$, while $\mathbf{s}_{b_{i,j}}$ and $\mathbf{s}_{f_{i,j}}$ are the numerical discretization of the bed-slope and friction source terms in $C_{i,j}$, respectively. From inspection of Equations (7) and (8), it is apparent that the evolution of the vector $\mathbf{u}_i$ is driven by the evolution of three degrees of freedom per cell.

In order to evaluate the right-hand side of Equation (8), the values of the variables $h$, $hU$ and $\zeta = h + z$ are needed in properly defined quadrature points into the FV and at interfaces. If $u$ is the generic variable to be reconstructed and $u_{i,j}$ is the cell-averaged value of $u$ in $C_{i,j}$, it is possible to build a function $u_i(x)$, which approximates $u$ in $S_i$, with the third order of accuracy, by means of a linear combination of three second-degree polynomials $c_{i,j}(x)$:

$$
u_i(x) = \sum_{j=1}^{3} c_{i,j}(x) u_{i,j}, \quad x \in [x_{i-1/2}, x_{i+1/2}].
$$

Of course, the polynomials $c_{i,j}(x)$ are chosen in order to preserve the averaged values of $u$ in the FV of $S_i$ (Wang 2002).

Once the conserved variables are reconstructed at the interfaces, fluxes can be evaluated. In general, at the interface between the SVs $S_i$ and $S_{i+1}$, two different reconstructions of the conserved variables are available, from the left and from the right, respectively: in this case, the fluxes are calculated by means of a homogeneous Riemann solver, approximate or exact. In contrast, the reconstruction is continuous through internal interfaces and fluxes can be calculated analytically, with savings in computational resources: of course, discontinuities of the conserved variables may arise also at internal interfaces of a SV, due to variable limitations or to the presence of wetting–drying fronts, and also in this case a Riemann solver is needed.

At the generic interface where the reconstruction is discontinuous, let $\mathbf{u}_{i,j+1/2}$ and $\mathbf{u}_{i,j-1/2}^+$ be the values of the vector $\mathbf{u}$ reconstructed from the left and at the right, respectively: the HLLE Riemann solver $f_{i,j+1/2}^l = f_{i,j+1/2}^R (\mathbf{u}_{i,j+1/2}^l, \mathbf{u}_{i,j-1/2}^+)$ is used to evaluate the numerical flux (Harten et al. 1983; Einfeldt 1988). This solver has been chosen because it preserves the non-negativity of the depth by interface (Bouchut 2004), but alternative depth-positivity-preserving numerical fluxes (Godunov, Rusanov and so on) can be used as well.

It can be verified (Wang 2002) that uniform partitions of the SVs cannot guarantee the stability and the convergence of the method: this can be obviously attributed to the Runge phenomenon, i.e. the oscillation of the polynomial reconstruction at the boundaries of the reconstruction interval. In contrast, stable and convergent results are obtained when the FV are clustered at the boundaries of the SVs. In this paper, the partition suggested by Van den Abeele et al. (2007b) is used: if $\xi \in [0;1]$ is a local coordinate in the

---

Figure 1 | Definition sketch of the third-order spectral volume discretization of the computational domain.
SV $S_i$, such that $x = x_{i-1/2} + (x_{i+1/2} - x_{i-1/2}) \xi$, then the external cell interfaces are set at $\xi = 0$ and 1, while the internal interfaces are set at $\xi = 0.21$ and 0.79, respectively. Nicely, we observe that, if all the spectral cells are partitioned in the same manner, the choice of the polynomials $c_{ij}(x)$ is unique, with savings in memory and computational time.

While first-order schemes are diffusive enough to damp undesired numerical oscillations, the high-order reconstruction of the conserved variables in the SVs $S_i$ can cause the formation of undershoots and overshoots at the cell interfaces: these oscillations eventually increase, destroying the numerical solution and destabilizing the algorithm. In this paper, the total variation of the reconstructed variables is limited in each FV using a procedure inspired by Cockburn & Shu (1998) and Wang (2002). Recalling the TVD minmod function:

$$m(a_1, a_2, \ldots, a_n) = \begin{cases} s \min_{1 \leq i \leq n} |a_i|, & s = \text{sgn}(a_1) = \text{sgn}(a_2) = \cdots = \text{sgn}(a_n) \\ 0, & \text{otherwise} \end{cases}$$

(10)

the TVB minmod function can be defined as

$$\hat{m}(a_1, a_2, \ldots, a_n) = \begin{cases} a_1, & \text{if } |a_1| \leq M_u \Delta x_{i+1/2} \\ m(a_1, a_2, \ldots, a_n), & \text{otherwise} \end{cases}$$

(11)

where $M_u > 0$ is properly chosen as constant. With reference to the generic FV $C_{ij}$, if one of the conditions

$$\begin{cases} u_{i+1/2} - u_{i-1/2} - \hat{m}(u_{i+1/2} - u_{i-1/2}) = 0 \\ u_{i+1/2} - u_{i-1/2} + \hat{m}(u_{i+1/2} - u_{i-1/2}) = 0 \end{cases}$$

(12)

is not verified, then the linear reconstruction $u_{ij}(x) = u_{ij} + \lambda(x - x_{ij})$ is assumed, where the following optimal formulation of the conserved variable slope is used (Berger et al. 2005):

$$\lambda_{ij} = m \left( 2 \frac{u_{i+1} - u_{i-1}}{\Delta x_{ij} + \Delta x_{i+1/2}}, 2 \frac{u_{i+1} - u_{i-1}}{\Delta x_{ij} + \Delta x_{i+1/2}} \right).$$

(13)

Discretization of bed-slope and friction source terms

In order to apply the bed-slope source term treatment by Audusse et al. (2004), the conserved variables $h$ and $hU$, and the water surface elevation $\zeta$, are reconstructed in each SV. After the reconstruction, the values of the bed elevation at the generic interface can be calculated as $\zeta_{ij+1/2} = \zeta_{ij+1/2} - h_{ij+1/2}$, and the modified values $\hat{h}$ of the water depth, which are actually used for flux and source term evaluation, are obtained by means of the following upwinding rule (‘hydrostatic reconstruction’):

$$\hat{h}_{ij+1/2} = \max(0, \ z_{ij+1/2} - \max(\zeta_{ij+1/2}, \ z_{ij+1/2})).$$

(14)

After calculating the unmodified interface fluid velocity values $U_{ij+1/2} = hU_{ij+1/2}/h_{ij+1/2}$, one has the modified vectors of the conserved variables at the interface

$$\hat{u}_{ij+1/2}^x = \begin{bmatrix} \hat{h}_{ij+1/2}^x \\ \hat{h}_{ij+1/2}^x U_{ij+1/2}^x \end{bmatrix}^T$$

(15)

which are those actually used to calculate the numerical flux $f_{ij+1/2} = f_{\text{HLLE}}(\hat{u}_{ij+1/2}, \hat{u}_{ij+1/2})$ at the interfaces between the FV. The numerical bed-slope source term $s_{b_{ij}}$ is calculated as

$$s_{b_{ij}} = \frac{g}{\Delta x_{ij}} \left[ 0 \begin{cases} \frac{(h_{ij+1/2}^x)^2 - (\hat{h}_{ij+1/2}^x)^2}{2} - \frac{\zeta_{ij+1/2}}{h_{ij+1/2}^x} \frac{\partial \zeta}{\partial x} dx \\ \hat{h}_{ij+1/2}^x \frac{\partial \zeta}{\partial x} dx \end{cases} \right].$$

(16)

The evaluation of the integral in Equation (16) can be easily accomplished by means of a proper quadrature formula, with the desired order of accuracy. Here, the Romberg rule can be used, obtaining

$$\int_{x_{ij+1/2}}^{x_{ij+1/2}^x} h \frac{\partial \zeta}{\partial x} dx = \frac{1}{6} [4(h_1 + h_2)(\zeta_2 - \zeta_1) + 4(h_2 + h_3)(\zeta_3 - \zeta_2) - (h_1 + h_3)(\zeta_3 - \zeta_1)]_{ij+1/2}$$

(17)

where the following positions have been made:

$$\begin{cases} h_1 = h_{ij+1/2}, \ h_2 = h_{ij}, \ h_3 = h_{ij+1/2} \\ \zeta_1 = \zeta_{ij+1/2}, \ \zeta_2 = \zeta_{ij}, \ \zeta_3 = \zeta_{ij+1/2} \end{cases}$$

(18)
and the subscript C stands for the quantities calculated at the centre of the FV, while the subscripts 1 and 2 refer to the quantities calculated at the left and at the right boundary, respectively. The numerical friction source term is

\[
s_{ij} = \frac{1}{\Delta x_{ij}} \left[ \int_{x_{ij-1/2}}^{x_{ij+1/2}} g h S_i \, dx \right].
\]  

(19)

and the integral in Equation (19) is approximated by means of Simpson’s quadrature rule:

\[
\int_{x_{ij-1/2}}^{x_{ij+1/2}} g h S_i \, dx = \frac{g \Delta x_{ij}}{6} \left[ (h S_i)_1 + 4(h S_i)_C + (h S_i)_2 \right].
\]  

(20)

### Advancing in time of the numerical solution

After the space discretization, Equation (8) can be regarded as a system of ordinary differential equations, and advancing in time can be accomplished by means of a suitable algorithm. If \( V \) is the vector of the unknown variables \( u_{ij} \), the system (8) can be succinctly rewritten as

\[
d\frac{V}{dt} = L(V)
\]

(21)

and its solution is approximated by means of the explicit third-order TVD Runge-Kutta scheme. We obtain the following sequence of sub-steps (Gottlieb et al. 2001):

\[
\begin{align*}
V^{(0)} &= V^n \\
V^{(1)} &= V^n \leftarrow V' = V^{(0)} + \Delta t L(V^{(0)}) \\
V^{(2)} &= \frac{3}{4} V^n + \frac{1}{4} V' \leftarrow V'' = V^{(1)} + \Delta t L(V^{(1)}) \\
V^{n+1} &= \frac{1}{3} V^n + \frac{2}{3} V'' \leftarrow V = V^{(2)} + \Delta t L(V^{(2)})
\end{align*}
\]

(22)

where the superscripts \( n \) and \( n + 1 \) refer to the time levels \( t_n \) and \( t_{n+1} = t_n + \Delta t \), respectively, while \( L \) represents the space discretization operator. Variable reconstruction and limitation is needed at each sub-step in order to evaluate the space discretization operator.

Of course, the direct application of this approach is not appropriate when friction source terms are present, because the time step restriction needed to ensure the algorithm’s stability is well below the ordinary CFL condition near wet-dry transitions (Murillo et al. 2009). This issue can be overcome, and the time step restriction can be relaxed, observing that the Runge-Kutta algorithm coincides with a convex combination of Euler sub-steps, which can in turn be treated in a semi-implicit fashion: at each sub-step the fluxes and the bed-slope source term are treated explicitly, while the friction source term is treated implicitly. We obtain for the generic Euler sub-step

\[
\begin{align*}
\begin{cases}
V^* &= V^{(0)} + \Delta t L_c(V^{(0)}) \\
V' &= \Delta t L_t(V) = V^*
\end{cases}
\end{align*}
\]

(23)

where the space operator \( L_c \) refers to fluxes and bed-slope source terms, \( L_t \) refers to friction source terms, the superscript \( s \) is related to the \( s \)th Euler sub-step of the Runge-Kutta scheme, while the asterisk refers to the quantities calculated after the explicit sub-step. In order to ensure the third-order accuracy in space of the implicit sub-step, we observe that at the generic abscissa \( x \) the implicit Euler sub-step is

\[
hU(x) + \Delta t g h^*(x) S_t(x) = hU^*(x).
\]

(24)

We then also apply reconstruction and limitation of the variables \( h^* \) and \( hU^* \) after the explicit sub-step, use the formula (24) at each quadrature point of the friction source integral (20) in the FV \( C_{ij} \) and finally obtain

\[
hU_{ij} = \frac{1}{6} [hU_1 + 4hU_C + hU_2]_{ij}.
\]

(25)

Since the implicit sub-step is unconditionally stable, the global time step is bounded only by the time step restriction defined by the CFL condition.

### The well-balanced scheme in the case of moving wet-dry transitions

The well-balanced SV model described in the preceding subsection cannot directly cope with the cases of moving wet-dry frontiers, and slight modifications have to be introduced.
in order to ensure the positivity of the scheme and the C-property when wet–dry transitions are present: these modifications are based on the fact that in SV schemes a sub-gridding property can be exploited, and we can switch from spectral cell variable reconstructions to local finite-volume reconstructions, where needed.

First, a wet–dry tolerance height $\varepsilon_h$, close to zero, is defined: if $h_{ij} > \varepsilon_h$, the finite-volume $C_{ij}$ is considered wet, otherwise it is dry. Then, we consider a spectral cell as wet only if all the FV contained are wet: if one or more, but not all, FV are wet, the spectral cell is said to be partially wet; if no FV is wet, the spectral cell is dry.

Finally, the following supplementary rules are applied:

- the fluxes between two dry FV are null by definition, and do not need to be calculated;
- the high-order variable reconstruction is accomplished only in wet spectral cells;
- if a wet FV is contained in a partially dry spectral cell, a linear reconstruction of the variables is adopted, using only the information available in surrounding wet FV;
- if the FV $C_{i,k}$ is wet and $C_{i-1,1}$ is dry, then variables are limited in the cell $C_{i,k}$; similarly, if the FV $C_{i,1}$ is wet and $C_{i-1,k}$ is dry, then variables are limited in the cell $C_{i,1}$;
- if the FV $C_{ij}$ is wet, while $C_{ij-1}$ and $C_{ij+1}$ are dry, no reconstruction is accomplished in the FV;
- the reconstruction of the variables being at most quadratic, the minimum value $h^\text{min}_{ij}$ of the depth $h$ is immediately known after the reconstruction: if $h^\text{min}_{ij}$ is less than $\varepsilon_h$, then $h$, $\zeta$ and $hU$ are limited (‘limitation by depth’), in order to ensure the depth positivity of the reconstruction in the whole FV;
- if momentum $hU$ is limited, then Equation (13) is used to linearly reconstruct the velocity $U$, instead of $hU$.

**Positivity and well-balanced properties of the proposed scheme**

We have observed that the use of a numerical model for the approximate solution of the shallow-water equations with wet–dry transitions is feasible for practical purposes only if there are sufficient guarantees about the stability of the scheme and the physical soundness of its results. In particular, the scheme is also required to satisfy the non-negativity of the solution in the case of wet–dry transitions and to satisfy special solutions such as the water-at-rest condition (C-property): models which are not proven to satisfy these requirements should always be rejected in engineering practice. The following Theorem 1 states the satisfaction of the C-property and supplies a guide for the choice of the time step length $\Delta t$ in order to preserve the depth-positivity of the solution.

**Theorem 1.** The third-order SV Scheme (8), for the approximate solution of the shallow-water Equation (1), satisfies the C-property exactly also in the case of emerging bottom, and preserves the non-negativity of the water depth $h$ if the time step length $\Delta t$ is chosen in order to satisfy the following CFL condition:

$$\frac{\Delta t}{\Delta x_{ij}} \sigma_{ij} \leq \frac{1}{12}, \quad i = 1, 2, \ldots, NS, \quad j = 1, 2, 3 \quad (26)$$

where $\sigma_{ij}$ is a superior bound of the signal speed in the FV $C_{ij}$.

The definition of $\sigma_{ij}$, and the proof of Theorem 1, given following the approach by Perthame & Shu (1996), are contained in the Appendix (see http://www.iwaponline.com/jh/014/035.pdf).

Despite the fact that all the calculations accomplished in the following section are obtained with constant time step $\Delta t$, the estimate (26) can be used in order to dynamically enlarge or reduce the computational time step during calculations in practical applications (Mynett 1999).

**NUMERICAL RESULTS**

In this section, the results for a number of numerical standard tests, taken from the current scientific literature, are presented in order to demonstrate the capability of the numerical method to satisfy the C-property also in the presence of emerging bottom, to verify the order of accuracy of the scheme and to confirm the ability of the scheme to solve problems with moving wet–dry frontier, also in the presence of friction. The gravity constant $g$ was set equal to 9.81 m/s$^2$.

All the numerical experiments were run on an x64-based workstation, with six Intel Xeon processors X5650 (2.67 GHz), 12.0 GB RAM and Windows 7 operating system. The numerical codes were written in Visual Basic 6, and run in executable form.
C-property

This test, inspired from that by Gallardo et al. (2007), is devoted to verify the C-property, when wet–dry transitions are included. In a channel, 10 m long, we consider a bell-shaped mound, 5 m high, whose bed elevation is defined by

\[ z_b(x) = 5e^{-0.4ix - 5^2}, \quad x \in [0, 10] \]  

and the initial water surface level is set to \( \zeta = 2.5 \) m (see Figure 2).

The Manning formula is used for the friction source term, with roughness coefficient \( n = 0.05 \). The numerical model has been run in double precision until \( t = 0.5 \) s, using \( NS = 67 \) uniform spectral cells, corresponding to 201 degrees of freedom, time step \( \Delta t = 0.0005 \) s and wet–dry tolerance \( \varepsilon_m = 10^{-7} \) m, while the values \( M_b = M_{bU} = M_\zeta = 50 \) of the TVB limiter have been used: periodic boundary conditions have been imposed, in order to constrain possible spurious waves into the domain, with magnification of errors.

The inspection of Table 1 shows that the stationary solution is preserved at the level close to the round-off error.

No negative depths appeared during the simulation, confirming that the procedure proposed for the reconstruction and limitation of variables satisfies the C-property and the positivity of the solution also in the presence of wet–dry transitions.

Accuracy test

This test, proposed by Xing & Shu (2006a), is used in order to verify the third-order accuracy of the scheme for smooth solutions. In a channel which is 1 m long, the following periodic bed elevation and initial conditions are assumed:

\[ z_b(x) = \sin^2(\pi x), \quad h(x, 0) = 5 + \cos(2\pi x), \]  

\[ hU(x, 0) = \sin(\cos(2\pi x)), \quad x \in [0, 1] \]  

(28)

together with periodic boundary conditions. For the friction source term, the Manning formula is used, with \( n = 0.05 \). The numerical simulation is stopped at \( t = 0.1 \) s, because later the solution becomes discontinuous and the evaluation of the accuracy for smooth solutions is not possible anymore.

The calculations have been repeated for an increasing number of uniform spectral cells, from 15 to 480, doubling \( NS \) and halving the time step \( \Delta t \), with \( \Delta t \) chosen to respect strictly the rule \( CFL = 1/12 \) (see Theorem 1 in this paper). The values \( M_b = M_{bU} = M_\zeta = 1,000 \) of the TVB limiter have been used in order to avoid excessive limitation of data in a problem where the second-order derivative of the solution is very high. Since there is no analytic solution for this problem, a reference simulation has been calculated at time \( t = 0.1 \) s on a very fine grid with \( NS = 7,680 \), and it has been used as the exact solution in order to compute the numerical errors. The results of calculations are summarized in Table 2: from inspection of this table, the third-order rate of convergence in the \( L_1 \) and \( L_\infty \) norms is clearly confirmed also in the presence of friction source terms. In Figure 3, the numerical solution obtained with \( NS = 30 \) SVs is compared with the solution obtained with the very fine grid: these solutions match perfectly with the solutions presented in the literature for the same test (Caleffi et al. 2006). It is interesting to observe how a modest number of SVs is needed to reach sufficient accuracy for this test case.
In order to explain the tendency in the orders of accuracy which are found in Table 2, we observe that, for increasing numbers of SVs, the order of accuracy tends to increase from values which are lower than the nominal ones, reaching the third order of accuracy in the so-called ‘asymptotic regime’ (which is attained for more than \( N_S = 120 \) spectral cells in this case). The effect of the increase of the order of accuracy for the refining grid is constantly observed in the literature (see, for example, Wang \( \text{et al.} \) 2002, tables II, III and IV). Of course, oscillations of the order of accuracy, especially in the case of complicated systems of differential equations, are also expected in the asymptotic regime. It remains to explain the slight drop in the order of accuracy which can be observed for \( N_S = 480 \) spectral cells. We observe that the test case used does not have an easily calculable analytical solution, and then errors were evaluated using a reference solution computed on a very fine grid, consisting of 7,680 spectral cells. The reference solution does not coincide with the analytical solution, of course: this means that, for \( N_S \) tending to infinity, the numerical solutions tend asymptotically to the analytical solution and not to the reference solution: this observation clarifies that, for values of \( N_S \) approaching 7,680, the use of an approximate reference solution for the evaluation of the errors has no meaning. In our case this effect is sensitive for \( N_S \) greater than 480.

In Table 2 are also reported (column labelled ‘Computational time’) the durations of the numerical experiments, expressed in milliseconds. Every spectral cell used for the SV3RK3 scheme contains three FV, and then a grid of \( N_S \) spectral cells consists of \( N_V = 3N_S \) degrees of freedom: the first column of Table 2 (‘\( N_V \)’) contains the number of degrees of freedom used during the calculations.

### Comparison with lower-order schemes

In order to compare the proposed SV numerical scheme with existing methods, a standard second-order finite-volume scheme, called FV2RK2 (see, for example, Audusse \textit{et al.} 2004), has been coded: the characteristics of this scheme are summarized in Table 3.

As in the preceding section, the second-order accuracy of the FV2RK2 scheme has been verified by means of the convergence test by Xing & Shu (2006a), repeating the calculations for an increasing number of FV \( N_V \), ranging from 45 to 5,760, doubling \( N_V \) and halving the time step \( \Delta t \), with \( \Delta t \) chosen to respect strictly the rule CFL = 1/4 (see Proposition 2.27 in Bouchut 2004). The results of
calculations are summarized in Table 4: from inspection of this table, the second-order rate of convergence in the $L_1$ and $L_\infty$ norms for the conserved variables $h$ and $hU$ is confirmed.

The SV method coincides with a finite-volume method, where reconstruction of variables is made by grouping the FV in clusters called spectral cells: a fair comparison between the SV3RK3 and FV2RK2 schemes can then be made, considering computational grids composed of the same number of FV. From inspection of Tables 2 and 4 it is apparent that, for a given number of FV, the SV3RK3 scheme is more accurate. In other words, we can say that the SV3RK3 scheme allows us to attain a given solution precision with a reduced number of computational degrees of freedom, if compared with lower-order finite-volume schemes: this confirms what is commonly found in the literature with reference to high-order schemes (Kreiss & Oliger 1971; Fletcher 1991). Of course, for a given number of FV, the SV3RK2 scheme is slower than the FV2RK2: actually, the gain of accuracy, obtained by considering more levels for the advancing in time, and more complicated formulae for the reconstruction and limitation of variables, does not come for free, and has a computational cost.

More interestingly, we can compare the numerical burden of the SV3RK3 and FV2RK2 schemes, for a given level of accuracy required: the $L_\infty$ norms of the error with reference to the specific discharge $hU$, contained in Tables 2 and 4, are represented as a function of the computational time in Figure 4. The plots representing the error for the SV3RK3 model (black triangles) and the error for the FV2RK2 model (white circles) cross at about $L_\infty = 2.5 \times 10^{-3}$ m$^2$/s: if the level of error required is lower than the error at the intersection point, the FV2RK2 model needs greater computational resources, and the same observation can be verified considering the other norms of the error. Of course, the point of intersection of the two error curves changes if we change the study case, or if we change the end time at which the calculations are stopped.

The results of the simulations with the SV3RK3 and FV2RK2 models confirm what is found in the literature, without exception: despite the fact that the solution error can be reduced by refining the computational grid, the grid refinement with higher-order accurate algorithms is computationally more efficient for sufficiently fine grids, that is for low error required (Abarbanel & Chertock 2000; Tkalich 2006). In the present case, a very short-duration one-dimensional test case ($t = 0.1$ s) has been sufficient to

<table>
<thead>
<tr>
<th>Variables reconstruction</th>
<th>Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction limitation</td>
<td>Minmod</td>
</tr>
<tr>
<td>Time advancing</td>
<td>TVD Runge Kutta 2</td>
</tr>
<tr>
<td>Numerical fluxes</td>
<td>HLLE</td>
</tr>
<tr>
<td>Geometric source term treatment</td>
<td>Hydrostatic reconstruction</td>
</tr>
<tr>
<td>Friction source term treatment</td>
<td>Implicit</td>
</tr>
<tr>
<td>CFL condition</td>
<td>1/4</td>
</tr>
</tbody>
</table>

More interestingly, we can compare the numerical burden of the SV3RK3 and FV2RK2 schemes, for a given level of accuracy required: the $L_1$ and $L_\infty$ norms of the error with reference to the specific discharge $hU$, contained in Tables 2 and 4, are represented as a function of the computational time in Figure 4. The plots representing the error for the SV3RK3 model (black triangles) and the error for the FV2RK2 model (white circles) cross at about $L_\infty = 2.5 \times 10^{-3}$ m$^2$/s: if the level of error required is lower than the error at the intersection point, the FV2RK2 model needs greater computational resources, and the same observation can be verified considering the other norms of the error. Of course, the point of intersection of the two error curves changes if we change the study case, or if we change the end time at which the calculations are stopped.

The results of the simulations with the SV3RK3 and FV2RK2 models confirm what is found in the literature, without exception: despite the fact that the solution error can be reduced by refining the computational grid, the grid refinement with higher-order accurate algorithms is computationally more efficient for sufficiently fine grids, that is for low error required (Abarbanel & Chertock 2000; Tkalich 2006). In the present case, a very short-duration one-dimensional test case ($t = 0.1$ s) has been sufficient to

<table>
<thead>
<tr>
<th>$L_\infty$</th>
<th>$L_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h$ (m)</td>
<td>$hU$ (m$^2$/s)</td>
</tr>
<tr>
<td>$h$ (m)</td>
<td>$hU$ (m$^2$/s)</td>
</tr>
<tr>
<td>NV</td>
<td>Error</td>
</tr>
<tr>
<td>45</td>
<td>0.0810</td>
</tr>
<tr>
<td>90</td>
<td>0.0350</td>
</tr>
<tr>
<td>180</td>
<td>0.0146</td>
</tr>
<tr>
<td>360</td>
<td>4.52 x 10$^{-5}$</td>
</tr>
<tr>
<td>720</td>
<td>7.74 x 10$^{-4}$</td>
</tr>
<tr>
<td>1,440</td>
<td>3.07 x 10$^{-4}$</td>
</tr>
<tr>
<td>2,880</td>
<td>7.67 x 10$^{-5}$</td>
</tr>
<tr>
<td>5,760</td>
<td>1.90 x 10$^{-5}$</td>
</tr>
</tbody>
</table>
recognize this effect: actually, when long duration computations are considered, diffusive errors of low-order schemes accumulate and the region of convenience of high-order schemes widens (Shu 2003).

Parabolic bowl

Sampson (2008) has proposed the analytical solution of the one-dimensional shallow-water equations for the case of unforced frictional flow in a parabolic channel, where the linear expression $S_l = g^{-1}U$ is assumed for the friction slope and $\tau$ is a roughness parameter. In a channel, 10,000 m long, the bed elevation is described by the quadratic function:

$$z_b(x) = h_0 \left( \frac{x}{a} \right)^2, \quad x \in [-5000, 5000]$$  \hspace{1cm} (29)

where $h_0 = 10$ m and $a = 3000$ m. The analytical solution for the water depth is defined by

$$h(x, t) = \max(0, \eta(x, t) - z_b(x)), \quad x \in [-5000, 5000]$$  \hspace{1cm} (30)

where

$$\eta(x, t) = h_0 - \frac{B^2 e^{-\tau t}}{4g} + \frac{a^2 B^2 e^{-\tau t}}{8g^2 h_0}$$

\[
\times \left( -\sigma \sin 2st + \left( \frac{r^2}{4} - s^2 \right) \cos 2st \right)
\]

\[
- xB e^{-(\tau/2)} \left( s \cos st + \frac{r}{2} \sin st \right)
\]

and the following definitions hold: $B = 5$ m/s, $\tau = 0.001$ s$^{-1}$, $p = \sqrt{8gh_0/a^2}$ and $s = \sqrt{p^2 - \tau^2}/2$.

Equations (30) and (31) express the fact that the mass of water oscillates, maintaining the free surface profile linear, tending exponentially to an asymptotic condition of rest: the initial conditions for the water depth $h$ are obtained by substitution of $t = 0$ in Equation (30), while the specific discharge $hU$ at time $t = 0$ is set to zero everywhere. For the numerical calculations with the SV3RK3 scheme, the computational domain has been subdivided into $NS = 100$ uniform spectral cells, corresponding to 300 degrees of freedom. The coefficients $M_h = M_{hU} = M_c = 50$ have been chosen for the simulation, and the solution has been advanced until $t = 6000$ s, with $\Delta t = 0.1$ s and $\varepsilon_n = 10^{-5}$ m.

In Figure 5, the results of the calculations at times $t = 1000, 2000, 3000, 4000, 5000$ and $6000$ s are compared with the analytical solution, with reference to the water surface elevation $\zeta$.

The agreement between numerical and analytical solutions for this very long simulation is good: in particular, the slowing down of the water mass due to the friction effects is correctly reproduced, together with the linear profile of the water surface, confirming the accuracy of the treatment of the geometric and frictional source terms. No negative water heights have been observed during the simulation and the total mass has been conserved up to round-off error, verifying the depth-positivity property of the algorithm, also in the presence of moving wetting–drying fronts, when a properly defined time step restriction is satisfied.

Steady transcritical flow with a shock over a hump

This test (Goutal & Maurel 1997; Alcrudo & Soares Frazão 1999) is widely used to evaluate the convergence of the numerical scheme to a steady condition, where the analytical solution is known (Caleffi et al. 2006; Xing & Shu 2006a, b; Caselles et al. 2009), and it is difficult due to the presence of a flow field discontinuity on a non-trivial bottom. In a channel, 25 m long, without friction, the bed elevation is described by

$$z_b(x) = \begin{cases} 
0.2 - 0.05(x - 10)^2, & \text{if } 8 < x < 12 \\
0, & \text{otherwise}
\end{cases}, \quad x \in [0, 25].$$  \hspace{1cm} (32)
For this test, the initial conditions are described by

\[ h(x, 0) = 0.5 - z_b(x), \quad hU(x, 0) = 0.36 h(x, 0), \quad x \in [0, 25]. \]  

(33)

The left boundary condition is defined by setting the water height \( h = 0.5 \) m and specific discharge \( hU = 0.18 \text{ m}^2/\text{s} \), while the right boundary condition is defined by setting the water depth \( h = 0.33 \) m. The analytical solution corresponding to the final steady condition can be found by coupling the Bernoulli theorem with the Rankine-Hugoniot condition.

In order to apply the SV3RK3 scheme, the computational domain has been subdivided into \( NS = 200 \) uniform spectral cells, corresponding to 600 degrees of freedom, while \( M_h = M_\zeta = 50 \) and \( M_{hU} = 0 \); the solution has been advanced until \( t = 300 \) s. In Figure 6, the results of the calculations at time \( t = 300 \) s are represented with reference to the water surface elevation \( \zeta \): it is apparent how the numerical model captures the steady analytical solution.
with great fidelity. In Figure 7, the numerical specific discharge is plotted at the time $t = 300$ s: the overall accuracy of the discharge evaluation is good, except at the hydraulic jump location. This issue is shared by many other numerical models (Caleffi et al. 2006) and is due to the shift between the theoretical position of the shock and the actual cell interface position.

**Dam break over a triangular sill**

In this test, the results of the numerical model are compared with the experimental data of the laboratory dam break (Hiver 2000; Soares-Frazão 2007). The experimental set-up consisted of a 38 m long horizontal channel, of rectangular cross section, with 0.75 m constant width, in which a gate divided a 15.5 m long reservoir from the dry bottom portion of the channel. A symmetric triangular-shaped sill, 6 m long and 0.40 m high, was located 10 m downstream of the gate. The right end of the channel consisted of a free outlet, and the Manning coefficient $n = 0.0125$ was evaluated for the channel bed. The initial conditions were characterized by a reservoir filled with 0.75 m of water: the sudden removal of the gate permitted us to reproduce the instantaneous failure of a dam. Seven gauges (G2, G4, G8, G10, G11, G13 and G20) were collocated at different distances from the gate (see Table 5) in order to measure the water depth variations during the transient consequent to the dam failure: in particular, gauges from G2–G11 were collocated at the left of the sill top, gauge G13 was collocated at the sill top while gauge G20 was collocated at the right of the sill.

After the gate removal at $t = 0$ s, the toe of the dam-break profile reached the sill at about $t = 3$ s, overtopped it at about $t = 4$ s and then invaded the dry zone on the right. Of course, the presence of the obstacle caused the partial reflection of the dam-break wave towards the closed end of the channel, where it was reflected again. Finally, the flow slowed down towards the rest, due to the dissipation of energy.

In order to perform this test, the computational domain is subdivided into $NS = 760$ uniform SVs, and the following parameters are used for the calculations: $\Delta t = 0.0001$ s, $e_h = 10^{-7}$ m, $M_h = M_{hu} = M_{\zeta} = 0$. In Figure 8, a comparison is made between numerical results (line) and experimental data (circles). It is apparent how the numerical model captures with fidelity the time of arrival of the first and subsequent reflected waves at all the gauges from G2 to G13. The water height is accurately predicted at gauges G2 and G4, together with the sudden increases of the depth due to the wave arrival, while it is slightly overpredicted at gauges from G8 to G13. The numerical results are less satisfactory only at gauge G20, where the water height is consistently

<table>
<thead>
<tr>
<th>Gauge</th>
<th>Abscissa (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G2</td>
<td>17.5</td>
</tr>
<tr>
<td>G4</td>
<td>19.5</td>
</tr>
<tr>
<td>G8</td>
<td>23.5</td>
</tr>
<tr>
<td>G10</td>
<td>25.5</td>
</tr>
<tr>
<td>G11</td>
<td>26.5</td>
</tr>
<tr>
<td>G13</td>
<td>28.5</td>
</tr>
<tr>
<td>G20</td>
<td>35.5</td>
</tr>
</tbody>
</table>
overpredicted, while the time of arrival of the transmitted waves are slightly retarded. The main discrepancies between the numerical solution proposed here and the experimental data are shared with other solutions proposed in the literature (Liang & Marche 2009): this shows that there are real water effects (turbulence, mixing with air, three-dimensional effects) which are not taken into account by the mathematical model of the shallow-water equations, and which can become important, especially downstream from obstacles.

CONCLUSIONS

Many numerical schemes have recently been presented, aiming at the high-order accurate solution of the shallow-water equations: these schemes, which are well balanced for water-at-rest states, or also for more general steady states, often neglect the problem of moving wet–dry frontiers and friction, which are fundamental for practical applications. In this paper, a third-order SV model for the approximate solution of the one-dimensional shallow water equations has been
described, and the desired properties of the scheme, namely the C-property and the depth-positivity, have been proven. In particular, it has been shown that, when a non-negativity-preserving numerical flux is used and the hydrostatic reconstruction of depths is adopted for the evaluation of geometrical source terms, the stability of the model can be ensured by a depth-positivity-preserving reconstruction of data and a properly defined CFL condition. This result shows that it is possible to construct robust high-order numerical schemes for the solution of the shallow-water equations when wetting–drying fronts are present, and supplies a guide for the choice of the time step during the calculations.

The numerical model proposed has been verified by means of numerous theoretical tests from the recent literature, which have confirmed the third order of accuracy, the exact satisfaction of water-at-rest conditions and the ability to manage wet–dry moving frontiers, also when non-trivial bathymetries and friction are present. Finally, a dam-break laboratory experiment has been reproduced for the choice of the time step during the calculations.

In future work, the theoretical and practical results obtained will be extended to approximate the solution of the De Saint Venant equations and the two-dimensional shallow-water equations, with moving wetting–drying frontiers, in order to face real-world applications.

ACKNOWLEDGEMENT

The writers wish to thank the three anonymous reviewers, who contributed to improve the paper.

REFERENCES


First received 25 March 2011; accepted in revised form 5 October 2011. Available online 30 January 2012.