World Scientific
www.worldscientific.com

Combining a Central Scheme with the Subtraction Method for Shallow Water Equations

Rony Touma ⓘ* and Elissa Malaeb ⓘ†

Computer Science and Mathematics
Lebanese American University, Byblos, Lebanon
*rony.touma@lau.edu.lb
†elissa.malaeb@lau.edu.lb

Christian Klingenberg ⓘ

Mathematics, University of Wuezburg
Wuezburg, Germany
klingenberg@mathematik.uni-wuerzburg.de

We present a new numerical scheme that is a well-balanced and second-order accurate for systems of shallow water equations (SWEs) with variable bathymetry. We extend in this paper the subtraction method (resulting in well-balancing) to the case of unstaggered central finite volume methods that computes the numerical solution on a single grid. In addition, the proposed scheme avoids solving Riemann problems occurring at cell boundaries as it employs intermediately a layer of ghost-staggered cells. The proposed numerical scheme is then implemented and validated. We successfully manage to solve classical SWE problems from the literature featuring steady states and other equilibria. The results of the study are consistent with previous research, which supports the use of the proposed method to solve SWEs.

Keywords: Shallow water equations; subtraction method; unstaggered central methods; well-balanced discretizations.

1. Introduction

The shallow water equations (SWEs) are a set of partial differential equations describing the flow of an in-compressible shallow layer of fluid with a free surface over a bottom topography under gravity. The bottom topography can either be flat or variable. SWEs are only admissible when the flow being studied is that

---

*Corresponding author.

of a shallow layer of fluid, i.e., when the fluid depth is much smaller than the wavelength of the waves being studied. They can model many natural phenomena such as tsunamis, tides, ocean currents and river floods.

Numerous numerical techniques have been employed to tackle shallow water problems with flat and varying bottom topographies. For instance, numerical techniques based on Riemann solvers were developed in Greenberg and LeRoux [1996] and Kim [2003], and recently, in Aleksyuk *et al.* [2022], an exact Riemann solver for the SWEs was developed. On another hand, well-balanced schemes were established in Touma [2009], Touma and Khankan [2012], Botta *et al.* [2004] and Greenberg and LeRoux [1996]. Additionally, authors in Touma [2016] and Audusse *et al.* [2004] addressed the drying of domains, such as Audusse *et al.* [2004] who used a hydrostatic reconstruction to obtain a well-balanced scheme that handles dry states ($h = 0$). There are also Bryson *et al.* [2011] whose authors developed central upwind schemes, as well as Shu and Xing [2005], Lu and Qiu [2011], Xing and Shu [2006] and Noelle *et al.* [2007] whose authors used finite difference and finite volume WENO schemes to tackle such problems. In Guo *et al.* [2023], the authors developed a new well-balanced finite volume CWENO scheme for the SWEs. Also, the authors of Ciallela *et al.* [2022] demonstrated a well-balanced finite volume WENO method combined with the modified Patankar Deferred Correction time integration method. Schemes based on relaxation models were presented in Delis and Katsaounis [2003, 2005]. And recently, in Gaburro *et al.* [2018], Desveaux *et al.* [2016], Castro and Pares [2020], Arpaia and Ricchiuto [2020] and Busto and Dumbser [2022], well-balanced schemes to solve the SWEs were presented. In Kent *et al.* [2023], the authors used an iterated semi-implicit time-stepping scheme along with a finite-volume transport method, and the cubed sphere grid, hence a mixed finite element discretization of the SWEs. Whereas in Kaptsov *et al.* [2022], the authors used conservative invariant finite difference schemes for the modified SWEs in Lagrangian coordinates.

On another hand, many schemes and methods have been employed recently on solving nonlinear PDE's, and due to which many useful results have been found such as in Gao *et al.* [2023], Chen *et al.* [2023a, 2023b], Zhao *et al.* [2022], Liu *et al.* [2023a, 2023b], Yin and Lu [2023], Yin *et al.* [2021, 2022], Lu *et al.* [2021], Lu and Chen [2021], Dong [2023], Xin *et al.* [2021], Buachart *et al.* [2014] and Sepehrirahnama *et al.* [2020].

As previously stated in Touma and Khankan [2012], the NT-scheme developed in Nessyahu and Tadmor [1990] is a non-oscillatory central scheme that approximates solutions to hyperbolic systems such as SWE systems. Its requirement for two staggered grids is a disadvantage, though. Jiang *et al.* [1998] created the first unstaggered outlook on NT-schemes. We can therefore think of the unstaggered central schemes (UCS) developed in Touma [2009] as being inspired by the previously mentioned schemes, with the advantages of avoiding Riemann solvers and staggered grids, working with only one original grid, a staggered ghost grid, and projecting back the numerical solution onto the original grid.

In this paper, we present a new numerical method for solving SWE problems in one dimension. It is second-order accurate, well-balanced, unstaggered, and uses the subtraction method. Its extra advantage is the ability to apprehend all types of steady states of the system at the discrete level.

The new method uses the same approach as the reconstruction technique previously employed in Kanbar *et al.* [2020]. Instead of resolving for the unknown solution $\mathbf{U}$, here, we will be evolving the error function $\mathbf{\Delta U}$ between $\mathbf{U}$ and a given steady state of the system $\mathbf{\tilde{U}}$, such that $\mathbf{\Delta U} = \mathbf{U} - \mathbf{\tilde{U}}$. The steady state we will use is the lake at rest steady state.

Our paper will be sorted as follows. In Sec. 2, we will present our UCS with the subtraction method. In Sec. 3, several common problems from the literature will be solved using the well-balanced UCS method presented in Touma and Khankan [2012] to compare their results to results obtained using UCS with subtraction. Some problems will also be compared to different methods found in the literature. And finally, we will have our concluding remarks in Sec. 4.

2. UCS with Subtraction Method

The SWEs system in one space dimension writes as

$$\begin{cases} \partial_t U + \partial_x f(U) = S(U), \\ U(x,0) = U_0(x), \end{cases} \tag{1}$$

where $x \in \mathcal{D} \subset]-\infty, \infty[$ and $t \in \mathbb{R}^*$. The unknown vector solution, the flux function and the source term in (1) are, respectively, as follows:

$$U = \begin{pmatrix} h \\ hu \end{pmatrix}, \quad f(U) = \begin{pmatrix} hu \\ hu^2 + \dfrac{1}{2}gh^2 \end{pmatrix}, \quad S(U) = \begin{pmatrix} 0 \\ -gh\dfrac{db}{dx} \end{pmatrix}. \tag{2}$$

In system (1) and (2), $h(x,t)$, $u(x,t)$, and $b(x)$ denote the height of the water column, the water velocity, and the topography (waterbed) function, respectively. The constant $g = 9.812$ denotes the gravitational constant.

The derivation involves two main physical principles, mainly the conservation of mass and the conservation of momentum.

By the definition of shallow waters, we assume that the wavelength is much bigger than its average height. This allows us to assume that the water velocity is only horizontal, where the vertical component of the velocity vector can be assumed to be zero. We can also assume that the horizontal component is the same throughout the fluid depth. Another assumption to be taken is that the density of the fluid ρ is constant. The fluid to be studied is in-compressible.

Finally, we can write the SWEs as follows:

$$\left\{ \begin{pmatrix} h \\ hu \end{pmatrix}_t + \begin{pmatrix} hu \\ hu^2 + g\dfrac{h^2}{2} \end{pmatrix}_x = \begin{pmatrix} 0 \\ -gh\dfrac{db}{dx} \end{pmatrix}. \tag{3}$$

Note that this system becomes homogeneous if we have a flat bottom topography, $b(x) = \text{constant}$.

Furthermore, the resulting homogeneous system of equations is hyperbolic, and the eigenvalues and eigenvectors are real and linearly independent.

We consider system (1) and (2) explicitly shown in system (3), and we assume that $\tilde{U}$ is the lake at rest stationary solution of system (1). Then, the error between the unknown solution U and this steady state $\tilde{U}$ is $\Delta U = U - \tilde{U}$, and so $U = \Delta U + \tilde{U}$.

Since $\tilde{U}$ is a stationary solution, then

$$\tilde{U}_t = 0 \quad \text{and so} \quad f(\tilde{U})_x = S(\tilde{U}). \tag{4}$$

Now, substitute U with $\Delta U + \tilde{U}$ in system (1) along with the fact that $\tilde{U}_t = 0$

$$\Delta U_t + f(\Delta U + \tilde{U})_x = S(\Delta U + \tilde{U}). \tag{5}$$

Performing a term by term subtraction in Eqs. (4) and (5), we obtain

$$\Delta U_t + [f(\Delta U + \tilde{U}) - f(\tilde{U})]_x = S(\Delta U + \tilde{U}) - S(\tilde{U}).$$

The right-hand side simplifies as follows:

$$S(\Delta U + \tilde{U}) - S(\tilde{U}) = \begin{pmatrix} 0 \\ -g(\Delta h + \tilde{h})\dfrac{db}{dx} \end{pmatrix} - \begin{pmatrix} 0 \\ -g\tilde{h}\dfrac{db}{dx} \end{pmatrix} = \begin{pmatrix} 0 \\ -g\Delta h \dfrac{db}{dx} \end{pmatrix}.$$

Accordingly, instead of solving system (1) now, we will be solving the following initial value problem:

$$\begin{cases} \partial_t \Delta U + \partial_x [f(\Delta U + \tilde{U}) - f(\tilde{U})] = S(\Delta U), & x \in \mathcal{D} \subset \mathbb{R}, \quad t > 0, \\ \Delta U(x, 0) = \Delta U_0(x). \end{cases} \tag{6}$$

We divide the computational domain $\mathcal{D}$ in system (6) into control cells $C_i = [x_{i-1/2}, x_{i+1/2}]$ and staggered dual cells $D_{i+1/2} = [x_i, x_{i+1}]$, as shown in Fig. 1. The control cells are centered at the nodes x_i and have length Δx. The staggered dual cells are centered at the nodes $x_{i+1/2} = x_i + \frac{\Delta x}{2}$ and also have length Δx. The time step is labeled as Δt and is used to increment the time variable following the rule $t^{n+1} = t^n + \Delta t$. Δt is obtained dynamically to ensure the stability of the numerical schemes.

Given the numerical solution U_i^n, an approximation to the exact solution $U(x_i, t^n)$, of system (1), and knowing *a priori* a steady state $\tilde{U}$ of system (1), we define the error $\Delta U_i^n = U_i^n - \tilde{U}_i$ as a piecewise linear function at time t^n over the cell C_i. Thus, the exact solution $\Delta U(x_i, t^n)$ will be approximated by ΔU_i^n. Our goal is then to generate the values of the numerical solution ΔU_i^{n+1} at t^{n+1} on the control cells C_i.
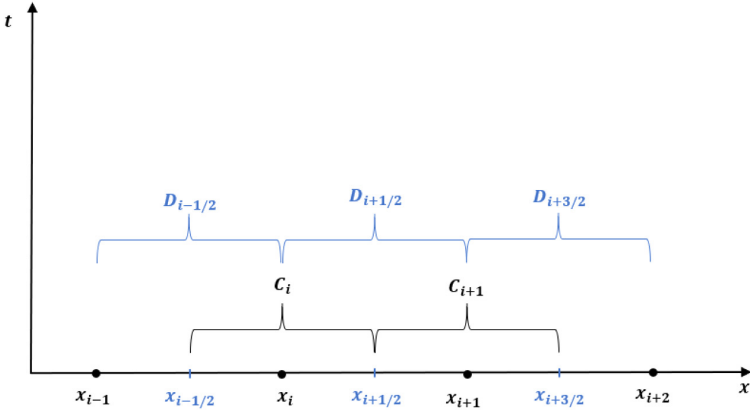
Fig. 1. Domain $\mathcal{D}$ partitioned into control cells $C_i = [x_{i-1/2}, x_{i+1/2}]$ and staggered dual cells $D_{i+1/2} = [x_i, x_{i+1}]$.

We define the piecewise linear function $\mathcal{P}_i(x,t)$ approximating the exact unknown error function $\Delta U(x,t)$ on the cells C_i

$$\Delta U_i^n = \frac{1}{\Delta x} \int_{C_i} \mathcal{P}_i(x,t) dx \approx \frac{1}{\Delta x} \int_{C_i} \Delta U_i(x,t) dx$$

and then

$$\mathcal{P}_i(x, t^n) = \Delta U_i^n + (x - x_i)(\Delta U_i^n)', \quad \forall\, x \in C_i,$$

where $(\Delta U_i^n)'$ is an estimate to the partial derivative $\frac{\partial \Delta U}{\partial x}(x_i, t^n)$ obtained by applying a limiting procedure of the numerical derivatives with the aid of the MC-limiter and given by

$$(\Delta U_i^n)' = \mathrm{MinMod}\left[\theta \frac{\Delta U_i^n - \Delta U_{i-1}^n}{\Delta x}, \frac{\Delta U_{i+1}^n - \Delta U_{i-1}^n}{2\Delta x}, \theta \frac{\Delta U_{i+1} - \Delta U_i^n}{\Delta x}\right],$$

where $\theta \in (1, 2)$, and the MinMod function is defined in Touma [2009].

From system (6), we get the following balance law:

$$\Delta U_t + [f(\Delta U + \tilde{U}) - f(\tilde{U})]_x = S(\Delta U). \tag{7}$$

Now, we integrate the balance law (7) over the domain $R_{i+1/2}^n = [x_i, x_{i+1}] \times [t^n, t^{n+1}]$

$$\iint_{R_{i+1/2}^n} \Delta U_t + [f(\Delta U + \tilde{U}) - f(\tilde{U})]_x \, dR = \iint_{R_{i+1/2}^n} S(\Delta U) dR.$$

We apply Green's formula to the double integral on the left-hand side. After a few simplifications, fixing and rearranging the integrals, we obtain

$$
\int_{x_i}^{x_{i+1}} -\Delta U(x, t^n)dx + \int_{t^n}^{t^{n+1}} [f((\Delta U + \tilde{U})(x_{i+1}, t)) - f((\tilde{U})(x_{i+1}, t))]dt
$$

$$
+ \int_{x_{i+1}}^{x_i} -\Delta U(x, t^{n+1})dx + \int_{t^{n+1}}^{t^n} [f((\Delta U + \tilde{U})(x_i, t)) - f((\tilde{U})(x_i, t))]dt
$$

$$
= \int_{t^n}^{t^{n+1}} \int_{x_i}^{x_{i+1}} S(\Delta U)\,dxdt. \tag{8}
$$

We estimate the integrals of the error function in the above equation by following the mean value theorem for the integral with respect to space. This can be indeed performed since $\Delta U(x, t)$ is supposed to be a piecewise linear function constructed at the center of the control cells. The integrals become

$$
\int_{x_i}^{x_{i+1}} -\Delta U(x, t^n)dx = -\Delta x \Delta U_{i+1/2}^n \quad \text{and}
$$

$$
\int_{x_{i+1}}^{x_i} -\Delta U(x, t^{n+1})dx = \Delta x \Delta U_{i+1/2}^{n+1}. \tag{9}
$$

Next we approximate the integrals of the flux function with second-order accuracy by applying the midpoint quadrature rule. This leads to

$$
\int_{t^n}^{t^{n+1}} [f((\Delta U + \tilde{U})(x_{i+1}, t)) - f((\tilde{U})(x_{i+1}, t))]dt
$$

$$
\approx \Delta t[f((\Delta U + \tilde{U})(x_{i+1}, t^{n+1/2})) - f((\tilde{U})(x_{i+1}, t^{n+1/2}))]
$$

and

$$
\int_{t^{n+1}}^{t^n} [f((\Delta U + \tilde{U})(x_i, t)) - f((\tilde{U})(x_i, t))]dt
$$

$$
\approx -\Delta t[f((\Delta U + \tilde{U})(x_i, t^{n+1/2})) - f((\tilde{U})(x_i, t^{n+1/2}))]. \tag{10}
$$

Plugging Eqs. (9) and (10) into Eq. (8), dividing by Δx, and rearranging, we obtain the forward evolution of ΔU_i^n onto the staggered grid at time t^{n+1}

$$
\Delta U_{i+1/2}^{n+1} = \Delta U_{i+1/2}^n - \frac{\Delta t}{\Delta x}[f(\Delta U_{i+1}^{n+1/2} + \tilde{U}_{i+1}) - f(\tilde{U}_{i+1})
$$

$$
- f(\Delta U_i^{n+1/2} + \tilde{U}_i) + f(\tilde{U}_i)] + \frac{1}{\Delta x} \int_{t^n}^{t^{n+1}} \int_{x_i}^{x_{i+1}} S(\Delta U)\,dxdt. \tag{11}
$$

The term $\Delta U_{i+1/2}^n$ occurring in the above equation is the forward projected value of ΔU_i^n onto the cells of the staggered grid. It is generated by expanding $\Delta U(x, t^n)$

in a Taylor series in space and then projecting the series onto the staggered grid

$$\Delta U_{i+1/2}^n = \frac{1}{2}(\Delta U_i^n + \Delta U_{i+1}^n) + \frac{\Delta x}{8}((\Delta U_i^n)' - (\Delta U_{i+1}^n)'),$$

where $(\Delta U_i^n)'$ is the numerical derivative estimating the spatial derivative of $\Delta U(x_i, t^n)$ calculated using a derivative limiting step.

The predicted values $\Delta U_i^{n+1/2}$ in Eq. (11) are calculated at the intermediate time $t^{n+1/2}$ by expanding $\Delta U(x, t^n)$ in a Taylor series in time and then using the first-order term

$$\Delta U(x_i, t^{n+1/2}) = \Delta U\left(x_i, t^n + \frac{\Delta t}{2}\right) \approx \Delta U(x_i, t^n) + \frac{\Delta t}{2}\Delta U_t(x_i, t^n).$$

Using the balance law (7), we obtain

$$\Delta U_i^{n+1/2} \approx \Delta U_i^n + \frac{\Delta t}{2}(-T_1 + T_2),$$

where the terms T_1 and T_2 are given by

$$T_1 = [f(\Delta U + \tilde{U}) - f(\tilde{U})]_x|_{(x_i, t^n)},$$
$$T_2 = [S(\Delta U)]|_{(x_i, t^n)}.$$

Hence we obtain

$$\Delta U_i^{n+1/2} = \Delta U_i^n + \frac{\Delta t}{2}[-(f_i^n)' + \tilde{f}_i' + S_i^n], \tag{12}$$

where $(f_i^n)'$ along with $\tilde{f}_i'$ denote the estimates to the derivatives in space at the point (x_i, t^n) of $f = f(\Delta U + \tilde{U})$ and $\tilde{f} = f(\tilde{U})$, respectively.

Note that $(f_i^n)'$ and $\tilde{f}_i'$ can be calculated using the MC-limiter or by using their respective Jacobians, such that

$$(f_i)' = \frac{\partial f}{\partial U} \cdot \frac{\partial U}{\partial x}\bigg|_i = J(f) \cdot (U_i)'.$$

Next we use the following discretization of the source term at time t^n:

$$S_i^n = S(\Delta U)|_{(x_i, t^n)} = \begin{pmatrix} 0 \\ -g\Delta h_i^n \left.\frac{\partial b}{\partial x}\right|_i \end{pmatrix}.$$

On the other hand, we estimate the integral on the right-hand side of Eq. (11) with second-order accuracy. For that we employ the trapezoidal quadrature rule along with centered differences, and the midpoint quadrature rule to obtain

$$\int_{t^n}^{t^{n+1}} \int_{x_i}^{x_{i+1}} S(\Delta U)dxdt \approx \Delta t \Delta x S(\Delta U_i^{n+1/2}, \Delta U_{i+1}^{n+1/2}).$$

The term to the right-hand side of the above equation is given by

$$S(\Delta U_i^{n+1/2}, \Delta U_{i+1}^{n+1/2}) = \begin{pmatrix} 0 \\ c_2 \end{pmatrix}, \tag{13}$$

with c_2 is given by $c_2 = -g(\frac{b_{i+1} - b_i}{\Delta x})\frac{\Delta h_{i+1}^{n+1/2} + \Delta h_i^{n+1/2}}{2}$.

Finally, the term ΔU_i^{n+1} denotes the backward projected value of the generated solution $\Delta U_{i+1/2}^{n+1}$ on the cells of the original mesh. It is calculated by expanding $\Delta U(x, t^{n+1})$ in a Taylor series in space and then projecting the series onto the original grid

$$\Delta U_i^{n+1} = \frac{1}{2}(\Delta U_{i-1/2}^{n+1} + \Delta U_{i+1/2}^{n+1}) + \frac{\Delta x}{8}((\Delta U_{i-1/2}^{n+1})' - (\Delta U_{i+1/2}^{n+1})'), \tag{14}$$

where $(\Delta U_{i+1/2}^{n+1})'$ is the limited numerical derivative of $\Delta U_{i+1/2}^{n+1}$ calculated using the MC-limiter.

So, now that we have ΔU_i^{n+1}, we can find the solution U_i^{n+1} at the cell-center of C_i at the next time step t^{n+1} as follows:

$$U_i^{n+1} = \Delta U_i^{n+1} + \tilde{U}_i.$$

We will summarize our scheme steps in order:

- Knowing the numerical solution U_i^n at time t^n and the lake at rest steady state $\tilde{U}_i$, we set $\Delta U_i^n = U_i^n - \tilde{U}_i$.
- Find the forward projected values

$$\Delta U_{i+1/2}^n = \frac{1}{2}(\Delta U_i^n + \Delta U_{i+1}^n) + \frac{\Delta x}{8}((\Delta U_i^n)' - (\Delta U_{i+1}^n)'),$$

where $(\Delta U_i^n)'$ is found using MC-limiter.
- Find the estimated values at the fractional time step $t^{n+1/2}$

$$\Delta U_i^{n+1/2} = \Delta U_i^n + \frac{\Delta t}{2}[-(f_i^n)' + \tilde{f}_i' + S_i^n],$$

where $(f_i^n)'$ and $\tilde{f}_i'$ denote estimates to the spatial partial derivatives of $f = f(\Delta U + \tilde{U})$ and $\tilde{f} = f(\tilde{U})$, respectively.
- Find the solution on the staggered dual cells at time t^{n+1}

$$\Delta U_{i+1/2}^{n+1} = \Delta U_{i+1/2}^n - \frac{\Delta t}{\Delta x}[f_{(\text{sum})_{i+1}}^{n+1/2} - \tilde{f}_{i+1} - f_{(\text{sum})_i}^{n+1/2} + \tilde{f}_i]$$

$$+ \Delta t S(\Delta U_i^{n+1/2}, \Delta U_{i+1}^{n+1/2}),$$

where

(1) $f_{(\text{sum})} = f(\Delta U + \tilde{U})$,
(2) $\tilde{f} = f(\tilde{U})$,
(3) $S(\Delta U_i^{n+1/2}, \Delta U_{i+1}^{n+1/2}) = \begin{pmatrix} 0 \\ c_2 \end{pmatrix}$.

- Find the solution needed on the original grid at time t^{n+1} by back-projecting $\Delta U_{i+1/2}^{n+1}$

$$\Delta U_i^{n+1} = \frac{1}{2}(\Delta U_{i-1/2}^{n+1} + \Delta U_{i+1/2}^{n+1}) + \frac{\Delta x}{8}((\Delta U_{i-1/2}^{n+1})' - (\Delta U_{i+1/2}^{n+1})'),$$

where $(\Delta U_{i+1/2}^{n+1})'$ is found using the MC-limiter.

Any numerical method used to solve the SWEs must have the C-property, which means that it must preserve rest. In other words, a lake at rest in the real world must also be at rest in the numerical model. This means that the steady state of a lake at rest must be preserved at the discrete level, i.e., $h + b =$ constant and $u = 0$. This new scheme we are developing, UCS with subtraction, satisfies the C-property because it preserves all steady states up to machine accuracy.

In fact, if U_i^n is a stationary solution of system (1), $U_i^n = \tilde{U}_i$, i.e., $\Delta U_i^n = 0$, we want to show that the updated numerical solution U_i^{n+1} at time t^{n+1} does not change over time, meaning that it is a stationary solution of the system of Eq. (1), i.e., $\Delta U_i^{n+1} = 0$. To show that, we must show that

(1) $\Delta U_i^{n+1/2} = 0$,
(2) $\Delta U_{i+1/2}^{n+1} = 0$,
(3) $\Delta U_i^{n+1} = 0$.

Proof.

(1) Using Eq. (12) and the fact that $\Delta U_i^n = 0$,

$$\Delta U_i^{n+1/2} = \Delta U_i^n + \frac{\Delta t}{2}[-f(\Delta U_i^n + \tilde{U}_i)' + f(\tilde{U}_i)' + S(\Delta U_i^n)]$$

$$= \frac{\Delta t}{2}[-f(\tilde{U}_i)' + f(\tilde{U}_i)']$$

$$= 0.$$

(2) Similarly, using Eqs. (11) and (13) and the fact that $\Delta U_i^{n+1/2} = 0$,

$$\Delta U_{i+1/2}^{n+1} = 0.$$

(3) Similarly, using Eq. (14) and the fact that $\Delta U_{i+1/2}^{n+1} = 0$,

$$\Delta U_i^{n+1} = 0. \qquad \square$$

3. Numerical Experiments

The partial differential equation in (7) can be written as

$$\Delta U_t + F(\Delta U)_x = S(\Delta U),$$

where $F(\Delta U) = f(\Delta U + \tilde{U}) - f(\tilde{U})$.

It was shown in Rogers *et al.* [2003] that the Jacobian matrix $J(F) = \frac{\partial F}{\partial \Delta U}$ is the same as the Jacobian matrix $J(f) = \frac{\partial f}{\partial U}$ corresponding to the PDE in system (1);

and this will be used in the CFL condition below: In all the following examples, the numerical results shown are made with a CFL $= 0.485$ such that

$$\Delta t = \text{CFL} \times \frac{\Delta x}{\alpha},$$

where $\alpha = \max_i |\lambda_i|$ where λ_i are the eigenvalues of the Jacobian matrix $J_i(f_i) = \frac{\partial f_i}{\partial U_i}$ over the cell C_i

$$\lambda_i = u_i \pm \sqrt{gh_i}.$$

In this section, we will be solving some problems common in the literature. We will mainly compare our computed solution with those previously obtained with a well-balanced UCS presented in Touma and Khankan [2012]. We will also compare our solutions to other results presented in the literature.

The steady state $\tilde{U}$ taken is that of a lake at rest

$$\tilde{U}_i^n = \begin{pmatrix} \tilde{h}_i^n \\ \tilde{h}_i^n \tilde{u}_i^n \end{pmatrix} \quad \text{such that } \tilde{h}_i^n + b_i = \text{constant and } \tilde{u}_i^n = 0.$$

We will use the following abbreviations: "IC" stands for initial condition, "WL" stands for water level $H = h + b$, u_0 stands for the initial water velocity $u(x, 0)$, h_0 stands for initial water height $h(x, 0)$, b stands for bottom topography $b(x)$, H_0 stands for initial water level $H_0 = h_0 + b$, t_f stands for final time, t_0 stands for initial time, "WB-UCS" stands for the well-balanced UCS, and "UCS-Sub" stands for the UCS with the subtraction method.

3.1. One-dimensional Toro's problem

We consider a simplified version of Toro's dam break problem, where the flow is only in the x-direction as defined in Touma and Khankan [2012], originally from Toro [2001].

The computational domain, which is the region of space that is being simulated, is divided into equal-sized segments using 200 grid points. The total length of the computational domain is 40 units.

The riverbed is flat ($b = 0$), $u_0 = 0$, $H_0 = h_0 + b = h_0$ where

$$h_0 = \begin{cases} 2.5, & 17.5 < x < 22.5, \\ 0.5, & \text{otherwise.} \end{cases}$$

The final time is $t_f = 1.4$.

Figure 2 shows the water level initially at time $t_0 = 0$. Figure 3 shows different snapshots of the water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line). Figure 4 shows the water level at the final time. We can observe in Figs. 3 and 4 an outward-propagating shock wave and an inward-propagating rarefaction wave. Both schemes lead to the same results at all times.
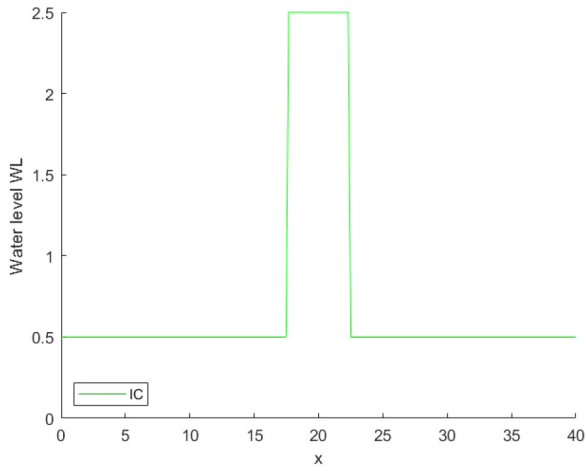
Fig. 2. One-dimensional Toro's problem: Initial water level at $t_0 = 0$.
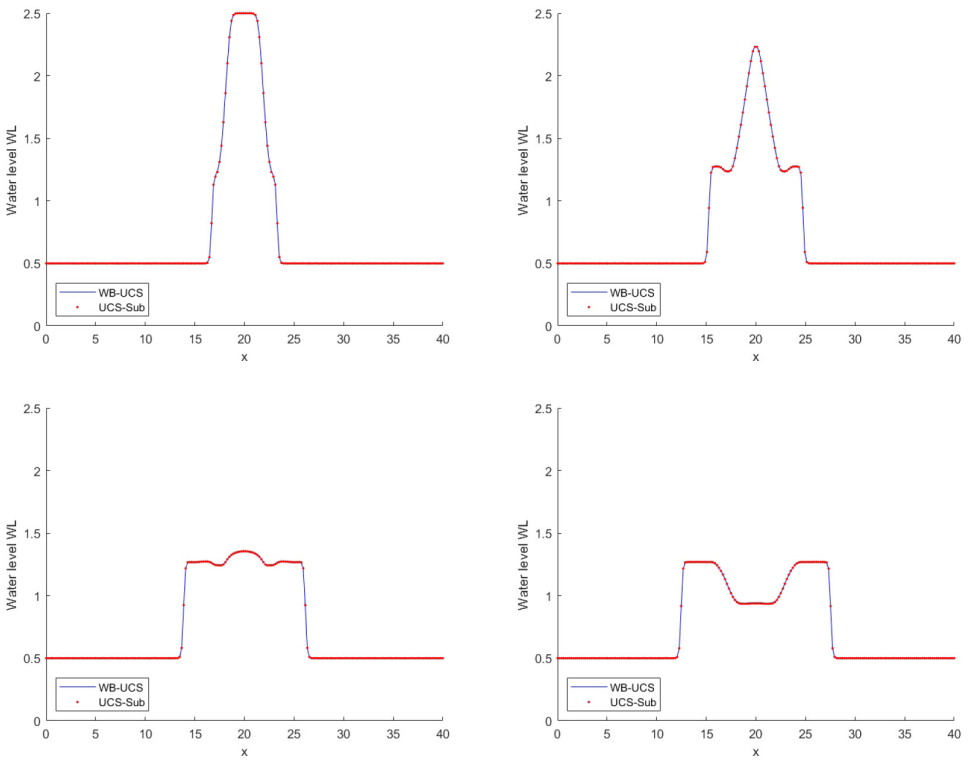

Fig. 3. Toro's problem: Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at different times before the final time.
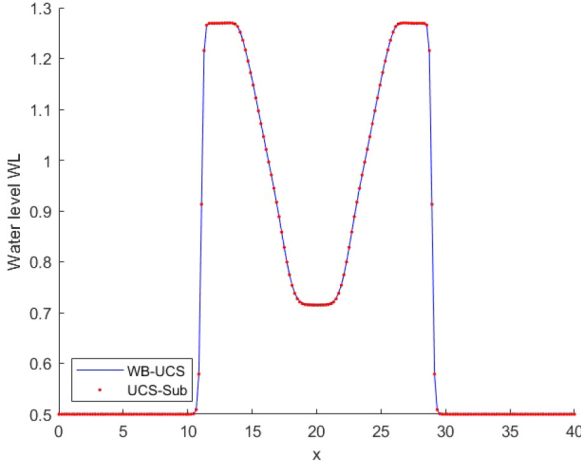
Fig. 4. Toro's problem: Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at $t_f = 1.4$.

3.2. Lake at rest test case

We consider this one-dimensional water flow over variable bottom topography problem. The computational domain is the range of x-values from 0 to 10. It is divided into 100 grid points. The waterbed function $b = b(x)$ is defined by

$$b = \begin{cases} 0.25, & 0 \leq x < 1, \\ g(x), & 1 \leq x \leq 9, \\ 0.30, & 9 < x \leq 10, \end{cases}$$

where

$$g(x) = (10e^{-x^2} + 15e^{-(x-2.5)^2} + 10e^{-(x-5)^2/2} + 6e^{-2(x-7.5)^2} + 16e^{-(x-10)^2})/20.$$

The initial conditions are $H_0 = 1$ and $u_0 = 0$. The final time is $t_f = 1$.

Figure 5 shows the water level initially at time $t_0 = 0$. Figure 6 shows the water level obtained at the end of the simulation, at time $t_f = 1$ using WB-UCS (solid line) and UCS-Sub (dotted line). The UCS-Sub method was able to produce a numerical solution that perfectly met the steady state requirement. This is a strong indication that the method is accurate and reliable. The numerical solution is in perfect agreement with the results obtained using the WB-UCS method and with the analytic equation solution.

3.3. Water flow test case over variable waterbed

We consider next the water flow test case on variable bathymetry as presented earlier in Touma [2009]. The computational domain is set to be the region on the
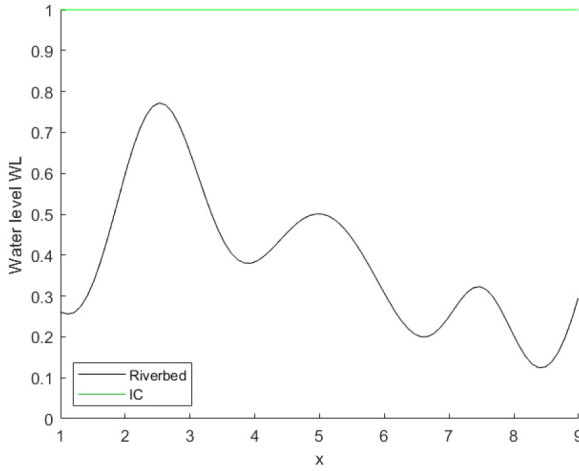
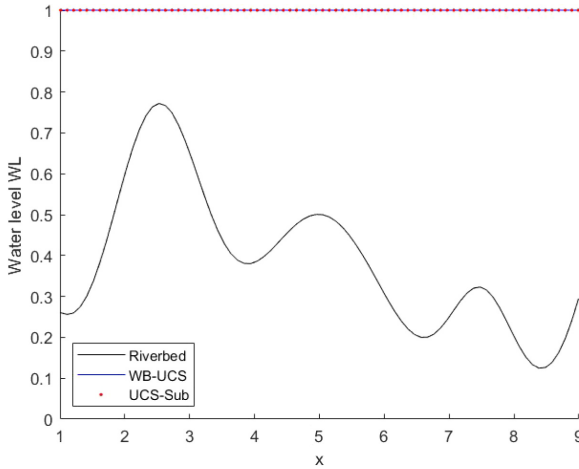Fig. 5. Lake at rest problem: Initial water level at $t_0 = 0$.

Fig. 6. Lake at rest problem: Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at $t_f = 1$.

x-axis satisfying $|x| \leq 10$ which we mesh using 2000 grid points. The waterbed bathymetry $b = b(x)$ is piecewise described by

$$b = \begin{cases} \dfrac{1}{5}\left(1 - \left(\dfrac{x}{2}\right)^2\right), & -2 \leq x \leq 2, \\ 0, & \text{otherwise.} \end{cases}$$

The initial water level $H_0 = 1$ and $u_0 = 1$. The final time is $t_f = 20$.

Figure 7 shows the water level initially at time $t_0 = 0$. Figures 8 and 9 show the water level obtained at different times and at the final time, respectively,
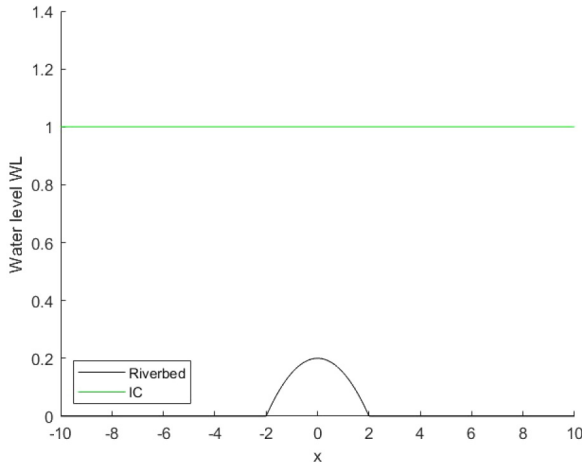
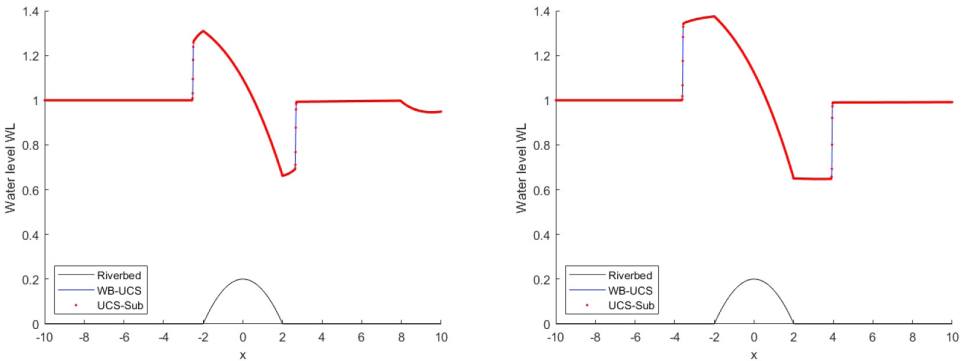Fig. 7. Water flow over variable bottom topography: Initial water level at $t_0 = 0$.

Fig. 8. Water flow over variable bottom topography problem: Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at different times before the final time.

using WB-UCS (solid line) and UCS-Sub (dotted line). The perfect alignment of results from the two schemes confirms the robustness and potential of the proposed UCS-Sub.

Figure 10 shows the water level at the final time obtained using WB-UCS and UCS-Sub in comparison with results obtained using NT+SB; the results were those presented in Touma [2009]. The results show a great match between all solutions.

3.4. *Dam break flow, no bottom topography*

We consider now the one-dimensional dam break problem previously presented in Delis and Katsaounis [2003]. Consider a channel of Length $L = 2000$. A dam is located at $x = 1000$. At the time $t_0 = 0$, the dam collapses. The goal is to find the
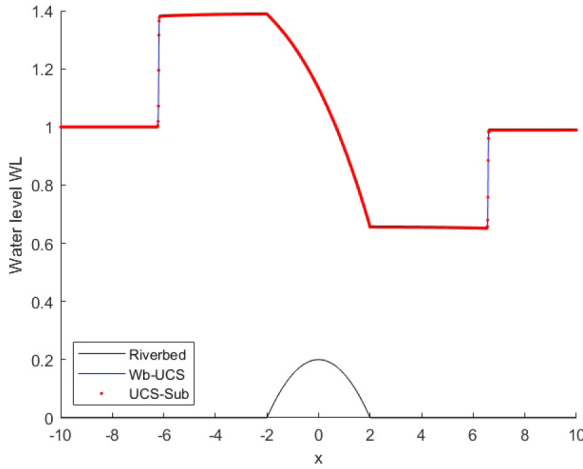
Fig. 9. Water flow over variable bottom topography problem: Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at $t_f = 20$.
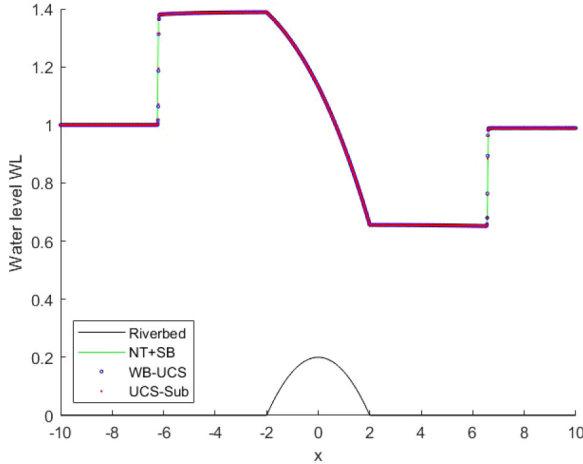

Fig. 10. Water flow over variable bottom topography: Water level obtained using WB-UCS and UCS-Sub compared with results obtained using the NT+SB method.

solution at time $t_f = 50$. The computational domain is the range of x-values from 0 to 2000. It is divided into 101 grid points to get $\Delta x = 20$.

The riverbed is flat $(b = 0)$, $u_0 = 0$, $H_0 = h_0 + b = h_0$ where

$$h_0 = \begin{cases} h_1, & x \leq 1000, \\ h_2, & x > 1000, \end{cases}$$

with $h_1 = 10 > h_2$. The solution to this problem features a shock wave propagating downstream and a rarefaction wave propagating upstream. The type of flow changes
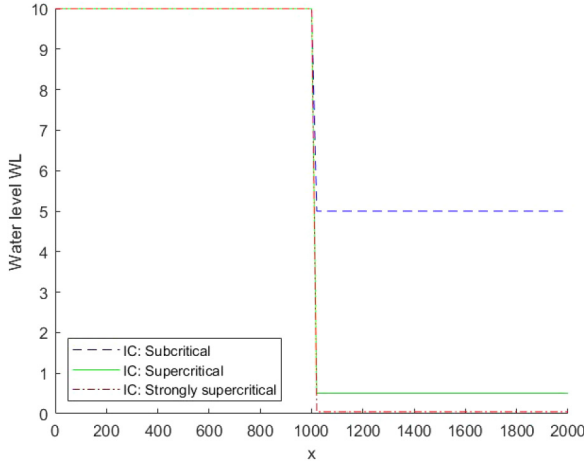
Fig. 11. Dam break flow: Initial water level at $t_0 = 0$. Subcritical stands for $h_0/h_1 = 0.5$, supercritical stands for $h_0/h_1 = 0.05$, and strongly supercritical sands for $h_0/h_1 = 0.005$.

between subcritical, supercritical, and strongly supercritical according to the water height downstream, as follows:

- if $h_2/h_1 \geq 0.5$: subcritical flow,
- if $h_2/h_1 < 0.5$: subcritical upstream, supercritical downstream,
- if $h_2/h_1 << 0.5$: strongly supercritical downstream.

Figure 11 shows the water level initially at time $t_0 = 0$. Figure 12 shows the water level obtained at the final time $t_f = 50$ using WB-UCS (solid line) and UCS-Sub (dotted line). The numerical results are consistent across all three cases: subcritical, supercritical, and strongly supercritical.

In Fig. 13, the initial conditions are adjusted and we compare our results with those shown in Kim [2003]. Now, the channel length is $L = 100$ and the dam is located at its center. Upstream water height is $h_1 = 1$ and downstream water height is $h_2 = 0.1$. Figure 13 shows the water level at time $t_f = 8$. On the left, we show the water level computed by WB-UCS and UCS-Sub as compared with the analytic solution [Kim (2003); Toro (2001)]. On the right, we show the water level computed by WB-UCS and UCS-Sub as compared with the solutions obtained using the HLLC approximate Riemann solver [Kim (2003)]. The precise results reveal excellent agreement across all of the solutions, confirming the potential of our scheme.

3.5. *Steady flow over a bump*

We consider now steady flow over a bump previously considered in Xing and Shu [2006]. This problem aims to study convergence in time toward steady flow over a bump. The numerical solution is computed on the region $x \in [0, 25]$ of the x-axis,
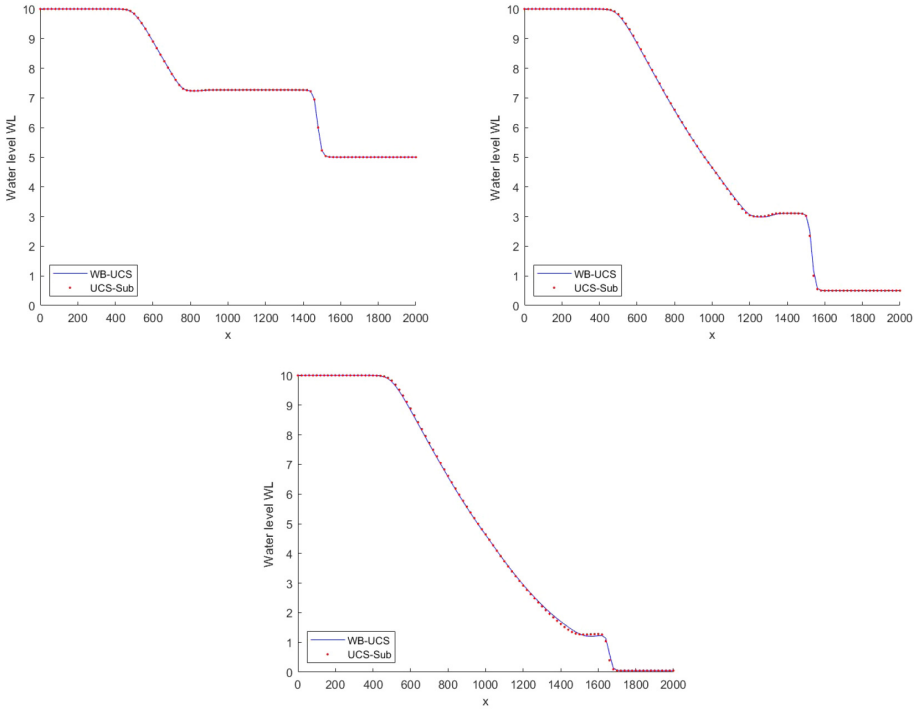
Fig. 12. Dam break flow problem: Subcritical, supercritical, strongly supercritical, respectively. Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at $t_f = 50$.
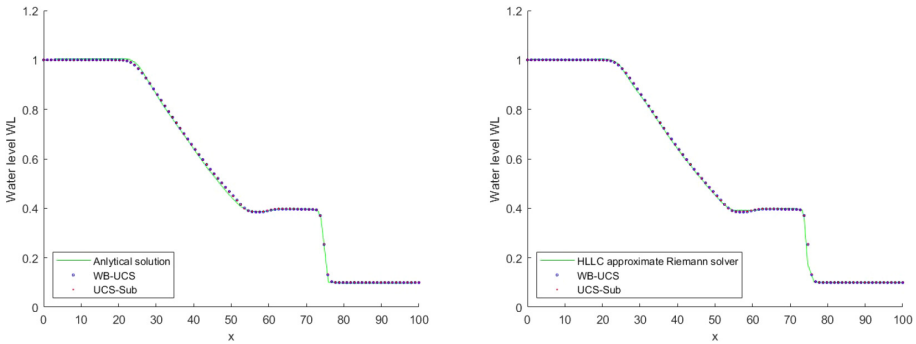

Fig. 13. Dam break flow: Water level obtained using WB-UCS and UCS-Sub compared with analytical solution (left) and HLLC approximate Riemann solver (right).

which we mesh using 201 grid points leading to $\Delta x = 0.125$. The bottom topography function $b = b(x)$ is given by

$$b = \begin{cases} \dfrac{1}{5} - \dfrac{1}{20}(x - 10)^2, & 8 \le x \le 12, \\ 0, & \text{otherwise.} \end{cases}$$

The initial conditions are $u_0 = 0$ for the initial velocity and $H_0 = H_d$ for the initial water level, where H_d is the water level downstream. The final time is $t_f = 200$. We can get three different cases: subcritical, transcritical without shock, and transcritical with shock. They vary with different boundary conditions.

- Subcritical flow: $(hu)_{\text{up}} = 4.42, H_d = 2$,
- Transcritical flow without shock: $(hu)_{\text{up}} = 1.53, H_d = 0.66(*)$,
- Transcritical flow with shock: $(hu)_{\text{up}} = 0.18, H_d = 0.33$,

where (hu) is the discharge, up stands for upstream, and (*): $H_d = 0.66$ is imposed only when the flow is subcritical.

Figure 14 shows the water level initially at time $t_0 = 0$. Figures 15–17 show the water level obtained in different cases using WB-UCS (solid line) and UCS-Sub (dotted line) at the final time $t_f = 200$. We have excellent superposing solutions, which demonstrate the validity of our approach. Table 1 shows the $L1$ error and order of convergence of steady flow over a bump problem in the subcritical case.

We see in Fig. 18 a comparison between the solutions obtained using WB-UCS and UCS-Sub compared with the analytic solutions Xing and Shu (2006) (left) and with the solutions obtained using WENO Xing and Shu (2006) (right). The figures show excellent agreement between the results which proves how promising our scheme is.

3.6. *Dam break over a rectangular bump*

We now study a one-dimensional problem with a fast-moving flow over an uneven riverbed as discussed in Touma and Khankan [2012]. The computational domain is [0,1500] and it is divided into equal-sized segments using 600 grid points. The numerical solution is resolved at the terminal time $t_f = 15$.
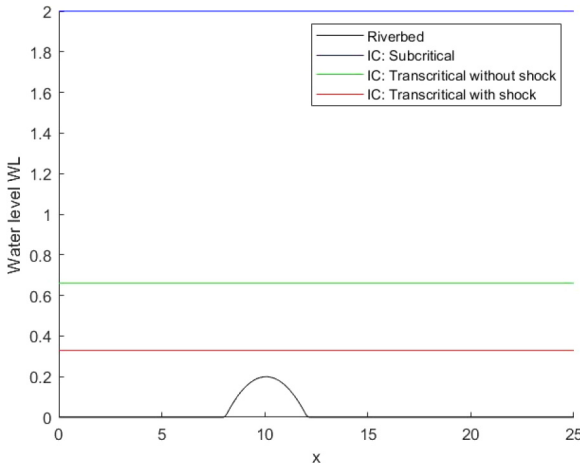

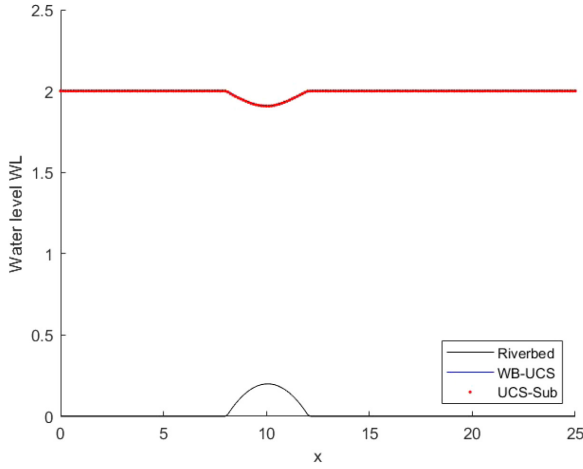Fig. 14. Steady flow over a bump: Initial water level at $t_0 = 0$.

Fig. 15. Steady flow over a bump: Case of a subcritical flow. Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at $t_f = 200$.
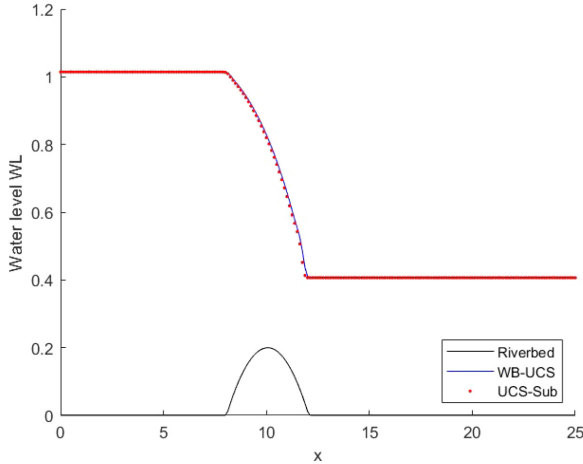

Fig. 16. Steady flow over a bump: Case of a transcritical flow without shock. Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at $t_f = 200$.

The bottom topography function $b = b(x)$ is given by

$$b = \begin{cases} 8, & |x - 750| < 375, \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

and the initial conditions are

$$H_0 = \begin{cases} 20, & x < 750, \\ 15, & \text{otherwise,} \end{cases} \tag{16}$$
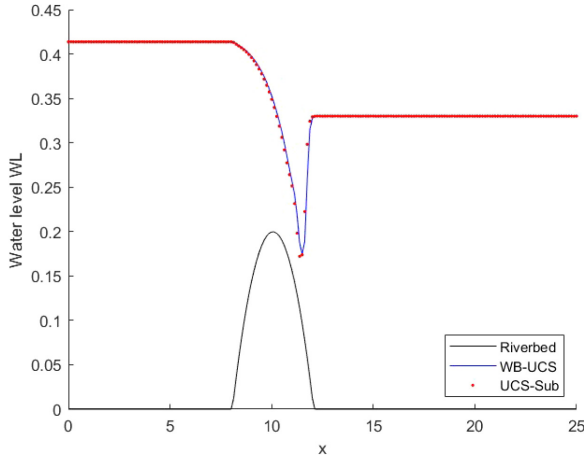
Fig. 17. Steady flow over a bump: Case of a transcritical flow with a shock wave. Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at $t_f = 200$.

Table 1. L_1 error and order of convergence of steady flow over a bump, subcritical case.

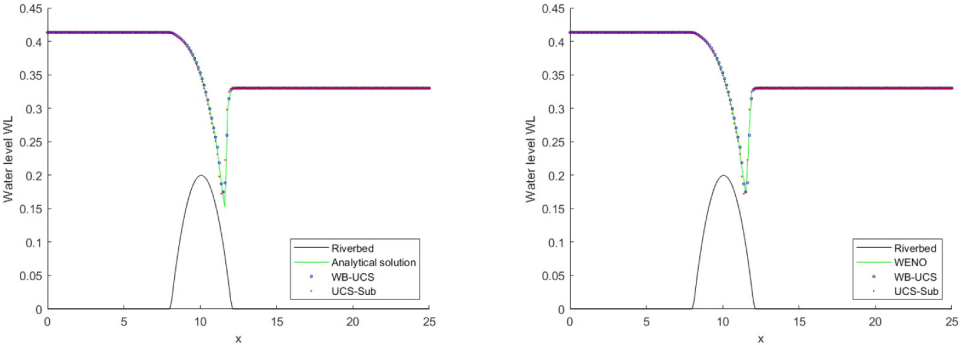| N	L_1 error h	L_1 order
200	$1.60e - 03$	
400	$3.85e - 04$	2.06
800	$7.73e - 05$	2.32

Fig. 18. Steady flow over a bump: Water level obtained using WB-UCS and UCS-Sub compared with the analytic solution (left) and the solution obtained using the WENO scheme (right).

for water level so we can deduce the water height function at $t_0 = 0$ and the velocity u_0 is set to be zero.

Figure 19 shows the water level initially at time $t_0 = 0$. Figures 20 and 21 show the water level obtained at different times and at the final time, respectively,
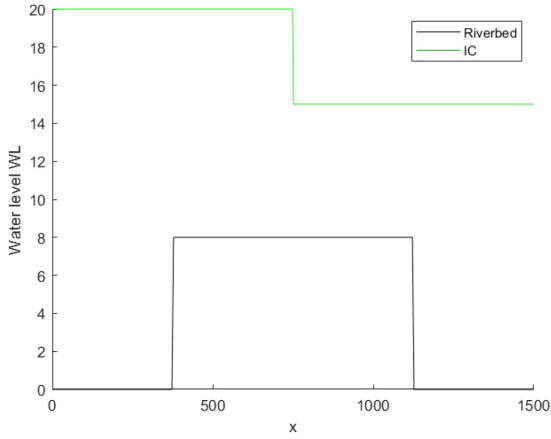
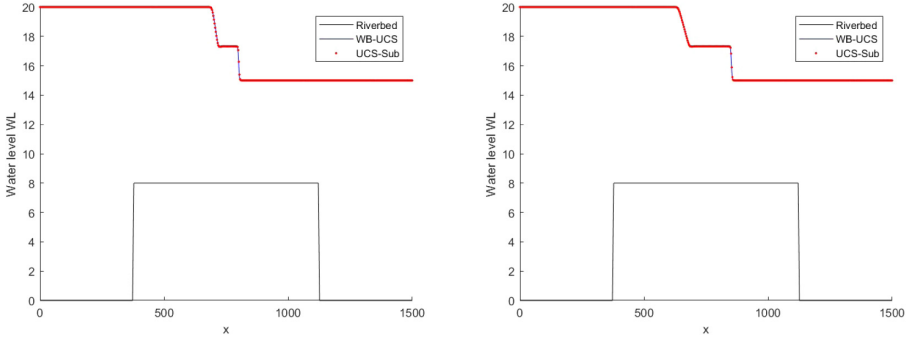Fig. 19. Dam break over a rectangular bump: Initial water level at $t_0 = 0$.

Fig. 20. Dam break over a rectangular bump: Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at different times before the final time.
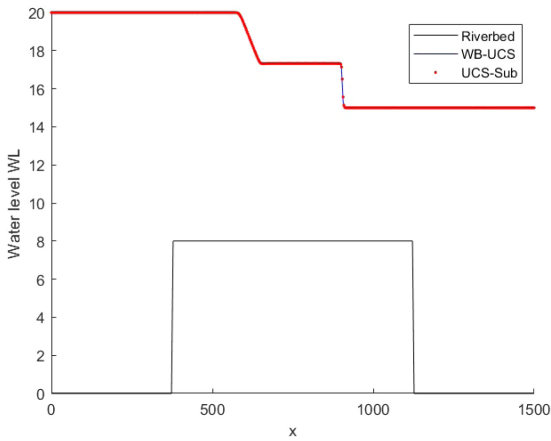

Fig. 21. Dam break over a rectangular bump: Water level obtained using WB-UCS (solid line) and UCS-Sub (dotted line) at $t_f = 15$.
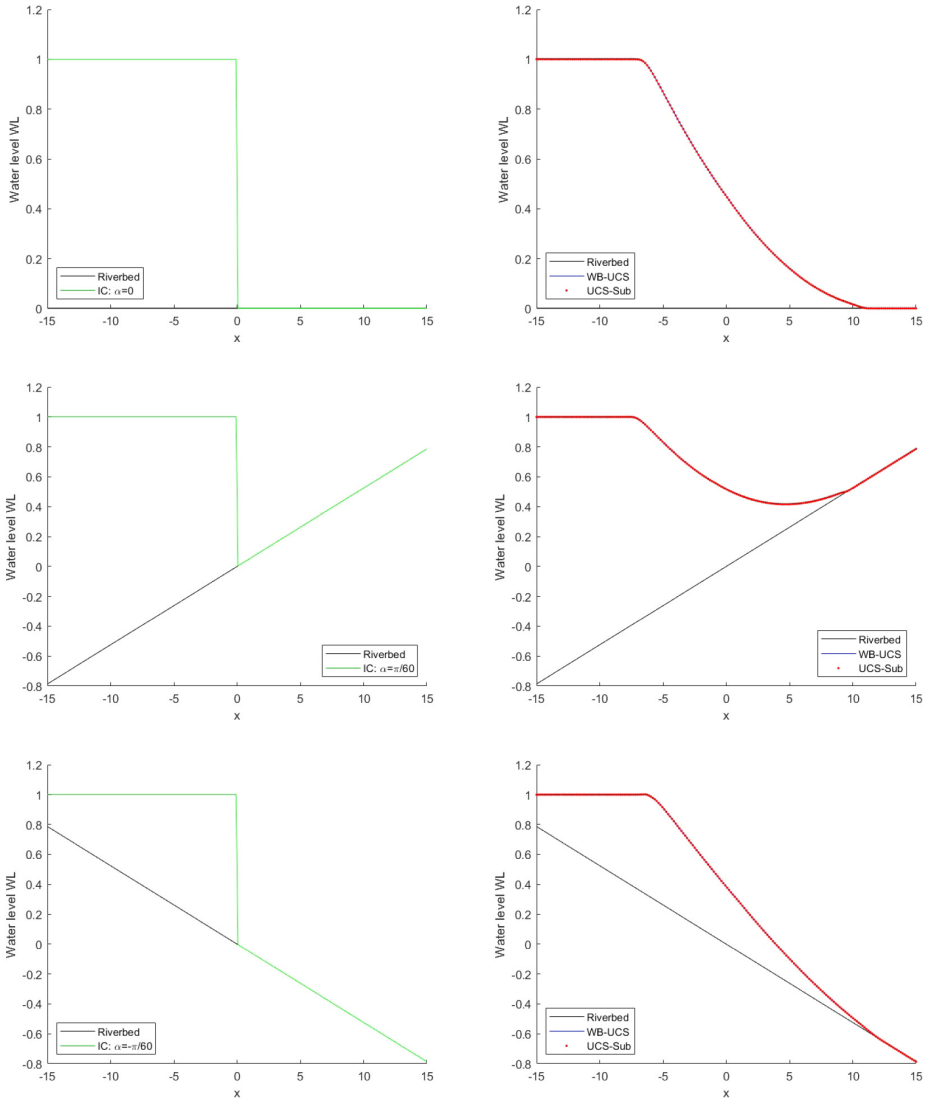
Fig. 22. Dam break problem over an inclined plane: Horizontally; the initial water level at time $t_0 = 0$ next to the water level obtained using WB-UCS + wet dry states (solid line) and UCS-Sub (dotted line) at $t_f = 2$.

using WB-UCS (solid line) and UCS-Sub (dotted line). The results show excellent agreement between both schemes.

3.7. *Dam break over an inclined plane*

We consider this one-dimensional problem featuring a dam break over inclined planes with different angles of inclination as presented in Touma [2016]. We mesh

the computational domain $[-15, 15]$ using 200 nodes and compute the numerical solution at the terminal time $t_f = 2$.

The waterbed function is defined by $b = b(x) = x\tan(\beta)$, where β represents the angle of inclination. The initial conditions are $u_0 = 0$ and

$$H_0 = \begin{cases} 1, & x < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

In Fig. 22, we see three rows; each row stands for a different angle of inclination. First row, $\beta = 0$, second row, $\beta = \pi/60$, and third row, $\beta = -\pi/60$. We see in each row the initial condition and next to it, the water level obtained at the final time $t_f = 2$ using WB-UCS with the wet and dry states (solid line) and UCS-Sub (dotted line). WB-UCS results were extracted from Touma [2016]. The results show excellent agreement between both schemes.

4. Conclusion

In this paper, we developed a new numerical method for solving the one-dimensional SWEs on variable topographies. The method is second-order accurate, unstaggered, and central, and it uses the subtraction approach to ensure well-balanced solutions. The approach has the benefits of UCS methods in avoiding Riemann solvers and staggered grid methods, as well as properly resolving any steady state of the considered system. The key concept is to evolve the error between the vector solution and any steady state of the system, rather than the unknown vector solution. We used the lake at rest stationary solution as the steady state in the error function in our work.

The problems we resolved produced very encouraging outcomes, and our numerical results are consistent with those found in other studies, thus validating the accuracy and the capability of the blended unstaggered central method with the subtraction method. In addition, the proposed numerical scheme showed a reduction in the computational time of about 15% as compared to the time required by other well-balanced techniques which usually necessitate building on special discretizations of the source term according to the discretization of the water height derivative, whereas UCS-Sub does not need to apply this as it has the C-property without any extra treatment. Not only does this make it faster as mentioned, but it also makes it easier to implement.

Generalizations of the presented one-dimensional technique to the two-dimensional context are currently being investigated. We also would like to test our scheme on different problems including friction. Even though we validated our scheme on a problem including a dry domain, but future work might check the treatment of the dry/wet states in all types of problems with discontinuous bottom topography and bowls for instance.

ORCID

Rony Touma ● https://orcid.org/0000-0003-0816-8724

Elissa Malaeb ● https://orcid.org/0009-0007-2966-1725

Christian Klingenberg ● https://orcid.org/0000-0003-2033-8204

References

Aleksyuk, A. I., Malakhov, M. A. and Belikov, V. V. [2022] "The exact Riemann solver for the shallow water equations with a discontinuous bottom," *J. Comput. Phys.* **450**, 110801.

Arpaia, L. and Ricchiuto, M. [2020] "Well-balanced residual distribution for the ALE spherical shallow water equations on moving adaptive meshes," *J. Comput. Phys.* **405**, 109173.

Audusse, E., Bouchut, F., Bristeau, M. O., Klein, R. and Perthame, B. [2004] "A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows," *SIAM J. Sci. Comput.* **25**(6), 2050–2065.

Botta, N., Klein, R., Langenberg, S. and Ltzenkirchen, S. [2004] "Well-balanced finite volume methods for nearly hydrostatic flows," *J. Comput. Phys.* **196**, 539–565.

Bryson, S., Epshteyn, Y., Kurganov, A. and Petrova, G. [2011] "Well-balanced positivity preserving central-upwind scheme on triangular grids for the saint-venant system," *ESAIM*: *Math. Model. Numer. Anal.* **45**, 423–446.

Buachart, C., Kanok-Nukulchai, W., Ortega, E. and Onate, E. [2014] "A shallow water model by finite point method," *Int. J. Comput. Methods* **11**(1), 1350047.

Busto, S. and Dumbser, M. [2022] "A staggered semi-implicit hybrid finite volume / finite element scheme for the shallow water equations at all Froude numbers," *Appl. Numer. Math.* **175**, 108–132.

Castro, M. and Pares, C. [2020] "Well-balanced high-order finite volume methods for systems of balance laws," *J. Sci. Comput.* **82**, 48.

Chen, S. J., Lu, X. and Yin, Y. H. [2023a] "Dynamic behaviors of the lump solutions and mixed solutions to a (2+1)-dimensional nonlinear model," *Commun. Theor. Phys.* **75**, 055005.

Chen, S. J., Yin, Y. H. and Lu, X. [2023b] "Elastic collision between one lump wave and multiple stripe waves of nonlinear evolution equations," *Commun. Nonlinear Sci. Numer. Simul.* **1**, 107205.

Ciallela, M., Micalizzi, L., Offner, P. and Torlo, D. [2022] "An arbitrary high order and positivity preserving method for the shallow water equations," *Comput. Fluids* **247**, 105630.

Delis, A. I. and Katsaounis, T. [2003] "Relaxation schemes for the shallow water equations," *Int. J. Numer. Methods Fluids* **41**, 695–719.

Delis, A. I. and Katsaounis, T. [2005] "Numerical solution of the two-dimensional shallow water equations by the application of relaxation methods," *Appl. Math. Model.* **29**, 754–783.

Desveaux, V., Zenk, M., Berthon, C. and Klingenberg, C. [2016] "Well-balanced schemes to capture non-explicit steady states: Ripa model," *Math. Comput.* **85**, 1575–1602.

Dong, J. [2023] "Well-balanced unstaggered central scheme based on the continuous approximation of the bottom topography," *Int. J. Comput. Methods* **20**(4) 2250058.

Gaburro, E., Castro, M. and Dumbser, M. [2018] "A well-balanced diffuse interface method for complex nonhydrostatic free surface flows," *Comput. Fluids* **175**, 180–198.

Gao, D., Lu, X. and Peng, M. S. [2023] "Study on the (2+1)-dimensional extension of Hietarinta equation: Soliton solutions and Bäcklund transformation," *Phys. Scr.* **98**, 095225.

Greenberg, J. M. and LeRoux, A. Y. [1996] "A well-balanced scheme for numerical processing of source terms in hyperbolic equations," *SIAM J. Numer. Anal.* **33**, 1–16.

Guo, W., Chen, Z., Qian, S., Li, G. and Niu, Q. [2023] "A new well-balanced finite volume CWENO scheme for shallow water equations over bottom topography," *Adv. Appl. Math. Mech.* **15**, 1515–1539.

Jiang, G. S., Levy, D., Lin, C. T., Osher, S. and Tadmor, E. [1998] "High-resolution non-oscillatory central schemes with non-staggered grids for hyperbolic conservation laws," *SIAM J. Numer. Anal.* **35**(6), 2147–2168.

Kanbar, F., Touma, R. and Klingenberg, C. [2020] "Well-balanced central schemes for the one and two-dimensional Euler systems with gravity," *Appl. Numer. Math.* **156**, 608–626.

Kaptsov, E. I., Dorodnitsyn, V. A. and Meleshko, S. V. [2022] "Conservative invariant finite-difference schemes for the modified shallow water equations in Lagrangian coordinates," *Stud. Appl. Mech.* **149**(3), 729–761.

Kim, D.-H. [2003] "Dam break flow analysis with approximate Riemann solver," *Water Eng. Res.* **4**(4) 175–185.

Kent, J., Melvin, T. and Wimmer, G. A. [2023] "A mixed finite-element discretisation of the shallow-water equations," *Geosci. Model Dev.* **16**, 1265–1276.

Liu, J., Wang, Y., Wang, L. Y., Yao, Q., Huang, C. C., Huang, H. Y. and Li, D. F. [2023a] "Several promising non-vdW multiferroic half-metallic nanosheets ACr2S4 (A=Li, Na, K, Rb): The first-principles researches," *Eur. Phys. J. Plus* **138**, 224.

Liu, K., Lu, X., Gao, F. and Zhang, J. [2023b] "Expectation-maximizing network reconstruction and most applicable network types based on binary time series data," *Physica D*: *Nonlinear Phenom.* **454**, 133834.

Lu, X. and Chen, S. J. [2021] "Interaction solutions to nonlinear partial differential equations via Hirota bilinear forms: One-lump-multi-stripe and one-lump-multi-soliton types," *Nonlinear Dyn.* **103**, 947–977.

Lu, X., Hui, H. W., Liu, F. F. and Bai, Y. L. [2021] "Stability and optimal control strategies for a novel epidemic model of COVID-19," *Nonlinear Dyn.* **106**, 1491–1507.

Lu, C. and Qui, J. [2011] "Simulations of shallow water equations with finite difference Lax-Wendroff weighted essentially non-oscillatory schemes," *J. Sci. Comput.* **47**(3), 281–302.

Nessyahu, H. and Tadmor, E. [1990] "Non-oscillatory central differencing for hyperbolic conservation laws," *J. Comput. Phys.* **87**(2), 408–463.

Noelle, S., Xing, Y. and Shu, C. W. [2007] "High order well-balanced finite volume WENO schemes for shallow water equations with moving water," *J. Comput. Phys.* **226**(1), 29–58.

Rogers, B. D., Borthwick, A. G. L. and Taylor, P. H. [2003] "Mathematical balancing of flux gradient and source terms prior to using Roe's approximate Riemann solver," *J. Comput. Phys.* **192**, 422–451.

Sepehrirahnama, S., Ong, E. T., Lee, H. P. and Lim, K. M. [2020] "Numerical modeling of free-surface wave effects on flexural vibration of floating structures," *Int. J. Comput. Methods* **17**(5), 1940016.

Shu, C. W. and Xing, Y. [2005] "High-order finite difference WENO schemes with the exact conservation property for the shallow water equations," *J. Comput. Phys.* **208**, 206–227.

Toro, E. F. [2001] *Shock-Capturing Methods for Free-Surface Shallow Flow* (Wiley, New York).

Touma, R. [2009] "Central unstaggered finite volume schemes for hyperbolic systems: Applications to unsteady shallow water equations," *Appl. Math. Comput.* **213**, 47–59.

Touma, R. and Khankan, S. [2012] "Well-balanced unstaggered central schemes for one and two-dimensional shallow water equation systems," *Appl. Math. Comput.* **218**, 5948–5960.

Touma, R. [2016] "Well-balanced central schemes for systems of shallow water equations with wet and dry states," *Appl. Math. Model.* **40**, 247–275.

Xin, X., Fengpeng B. and Kefeng L. [2021] "Numerical simulating open-channel flows with regular and irregular cross-section shapes based on finite volume Godunov-type scheme," *Int. J. Comput. Methods* **18**(4), 2050047.

Xing, Y. and Shu, C.-W. [2006] "High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms," *J. Comput. Phys.* **214**, 567–598.

Yin, Y. H. and Lu, X. [2023] "Dynamic analysis on optical pulses via modified PINNs: Soliton solutions, rogue waves and parameter discovery of the CQ-NLSE," *Commun. Nonlinear Sci. Numer. Simul.* **126**, 107441.

Yin, Y. H., Lu, X. and Ma, W. X. [2022] "Backlund transformation, exact solutions and diverse interaction phenomena to a (3+1)-dimensional nonlinear evolution equation," *Nonlinear Dyn.* **108**, 4181–4194.

Yin, M. Z., Zhu, Q. W. and Lu, X. [2021] "Parameter estimation of the incubation period of COVID-19 based on the doubly interval-censored data model," *Nonlinear Dyn.* **106**, 1347–1358.

Zhao, Y. W., Xia, J. W. and Lu, X. [2022] "The variable separation solution, fractal and chaos in an extended coupled (2+1)-dimensional Burgers system," *Nonlinear Dyn.* **108**, 4195–4205.