Accuracy aspects of conventional discretization methods for scalar transport with nonzero divergence velocity field arising from the energy balance equation

Zijlema, Marcel

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

RESEARCH ARTICLE

WILEY

Accuracy aspects of conventional discretization methods for scalar transport with nonzero divergence velocity field arising from the energy balance equation

Marcel Zijlema

Department of Hydraulic Engineering, Delft University of Technology, Delft, The Netherlands

Correspondence

Marcel Zijlema, Department of Hydraulic Engineering, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands.
Email: m.zijlema@tudelft.nl

Abstract

We are concerned with the numerical solution of a linear transport problem with nonzero divergence velocity field that originates from the spectral energy balance equation describing the evolution of wind waves and swells in coastal seas. The discretization error of the commonly used first-order upwind finite difference and first-order vertex-centered upwind finite volume schemes in one space dimension is analyzed. Smoothness of nondivergent velocity field plays a crucial role in this. No such analysis has been attempted to date for such problems. The two schemes studied differ in the manner in which they treat the scalar flux numerically. The finite difference variant is shock captured, whereas the vertex-centered finite volume approximation employs an arithmetic mean of the velocity and appears not to be flux conservative. The methods are subsequently extended to two dimensions on triangular meshes. Numerical experiments are provided to verify the convergence analysis. The main finding is that the finite difference scheme displays optimal rates of convergence and offers higher accuracy over the finite volume scheme, regardless the regularity of the velocity field. The latter scheme notably yields convergence rates of 0.5 and 0 in L^2-norm and L^∞-norm, respectively, when the velocity field is not smooth. A test case illustrating wave shoaling and refraction over submerged shoals is also presented and demonstrates the practical importance of flux conservation.

KEYWORDS

energy balance equation, finite difference method, finite volume method, flux conservation, nonzero divergence velocity, scalar transport equation

1 | INTRODUCTION

This article is devoted to the numerical approximation of solutions to the following linear transport equation written in the differential flux form

$$\frac{\partial F}{\partial t} + \nabla \cdot (\boldsymbol{v}\, F) = S, \quad \boldsymbol{x} = (x, y) \in \Omega \subset \mathbb{R}^2, \quad t > 0 \tag{1}$$

with suitable initial and boundary conditions. Here, $\boldsymbol{v}(\boldsymbol{x}, t) = (u, v)$ is the velocity field with the components u and v along the x and y coordinates, respectively, and $S(\boldsymbol{x}, t)$ is the nonconservative source term. In particular, we consider a class of problems under the assumption of a nonzero divergence velocity field, that is,

$$\nabla \cdot \boldsymbol{v} \neq 0.$$

Equation (1) is motivated by an application of spectral wave models. The equation under consideration is a special case of the so-called energy balance equation and is readily extended to integrate spectral distribution of wave energy over frequencies f and directions θ. In this respect, $F(f, \theta)$ represents the spectral density of energy at each location $\boldsymbol{x}$ and time t. The energy balance equation describes the time evolution of the energy density spectrum at a large number of locations in open ocean and coastal waters, when following a wave group (i.e., propagating with the group velocity $\boldsymbol{v}$ along a wave ray), while interacting with wind, current, and bottom. The processes of wave growth, redistribution and dissipation emerging from these interactions are known for being a complex physical and numerical problem. The concept of the energy balance equation is discussed in a didactic fashion in Reference 1. The spectral energy balance has been studied in detail in a series of classic papers,[2-4] and is of great interest for practical applications concerning the simulation of wind-generated waves and swells in coastal and naval engineering, oceanography and meteorology.[5]

Equation (1) poses a number of conceptual and numerical implications. In the case of pure advection, Equation (1) can be written as follows

$$\frac{\partial F}{\partial t} + \boldsymbol{v} \cdot \nabla F = S$$

with the left-hand side representing the Lagrangian rate of change of F along a path given by $d\boldsymbol{x}/dt = \boldsymbol{v}$. Here, the velocity field is divergence free. The method of characteristics and semi-Lagrangian schemes are suitable for such advection problems (see, e.g., Reference 1). On the other hand, in the case of a velocity field with nonzero divergence, the transport equation is no longer equivalent to the pure advection equation, and an additional source term must be provided

$$\frac{\partial F}{\partial t} + \boldsymbol{v} \cdot \nabla F = -(\nabla \cdot \boldsymbol{v})\, F + S.$$

The first term on the right-hand side is either a sink or a source of F at a given point, depending on the sign of $\nabla \cdot \boldsymbol{v}$ (divergence and convergence, respectively). Hence, F is not only propagated along a trajectory, but also created or destroyed locally.

Equation (1) is essentially a *local* balance where the divergence of flux $\boldsymbol{v} F$ is equally matched by the net source of F at each location under a steady-state condition. In the absence of sources and sinks, $\boldsymbol{v} F$ is thus a solenoidal vector field, that is, it is source free in every point $\boldsymbol{x}$ of Ω. This is also true for nonstationary cases with sources and sinks included. Because of the truly local character of Equation (1), this key feature should therefore be the primary condition for the design of numerical schemes, so as to maintain local conservation of flux $\boldsymbol{v} F$ at the discrete level. This can potentially reduce the solution error.

The abovementioned papers have reported numerical results for the purpose of verification and validation of the developed spectral wave models. To our knowledge, no analysis of convergence of the underlying numerical methods has been conducted in the literature. This motivates the research of the present study. The objective is to establish convergence rates of semidiscrete schemes frequently used in the spectral wave models through a truncation error analysis. Smoothness of nondivergent field $\boldsymbol{v}(\boldsymbol{x})$ plays a key role. The only study that we are aware of that takes into account the irregularity of the velocity field is due to Boyer,[6] where the convergence analysis of first-order implicit upwind finite volume schemes for linear transport problems on unstructured meshes is addressed. That paper proves convergence in $L^p(\Omega)$ with $p < +\infty$, though no error estimates are given.

The present article is organized as follows. The one-dimensional linear transport equation with nonzero divergence velocity and two commonly employed semidiscrete schemes are treated in Section 2 for the purpose of convergence analysis. The first scheme is obtained by applying the finite difference method and the second one is due to the vertex-centered finite volume approximation. Both schemes utilize a first-order upwind scheme for the approximation of the scalar flux. Next, in Section 3, the adopted semidiscrete methods are extended to unstructured grids of triangular cells in two dimensions. Numerical experiments demonstrating the behavior of the discretization error are presented in Section 4. Section 5

concerns a practical application arising from the spectral wave modeling to show the accuracy and efficiency of the discussed discretization schemes. Some concluding remarks are given in Section 6.

2 | METHODS AND ANALYSIS

The analysis will be carried out to the following initial-boundary value problem (IBVP)

$$\frac{\partial F}{\partial t} + \frac{\partial uF}{\partial x} = 0, \quad x \in \Omega \subset \mathbb{R}, \quad t > 0 \tag{2}$$

supplemented with proper and smooth initial and boundary conditions. In Equation (2), $F(x, t)$ is the sought solution and $u(x)$ is the given nondivergent velocity, that is, $\partial u/\partial x \neq 0$. Note that Equation (2) is linear and hyperbolic. We restrict ourselves in this article to examine the accuracy of spatial discretization of Equation (2). A few words on temporal discretization will be given later. Also it is not the purpose of this work to treat high-resolution schemes.

An important consideration of the analysis is that velocity $u(x)$ is bounded on Ω but may not be smooth over a subinterval of Ω. This suggests that u exhibits an abrupt change leading to the formation of a shock in F. As a final remark, the boundary conditions are not considered essential for the analysis as they are well defined and sufficiently smooth, and consequently their impact on the accuracy is rather limited. Therefore, the discussion on the boundary conditions is not included in the present study.

2.1 | The discrete formulation

The bounded domain Ω is discretized with a regular grid. It is assumed that the grid is sufficiently smooth, that is, quasi-uniform with the so-called algebraic stretching in the sense of Turkel.[7] The grid points are located at $\{x_m | m = 0, \dots, M\}$ and the grid size is given by $\Delta x_m = x_m - x_{m-1}$ ($m = 1, \dots, M$). The discrete solution $F_m(t) \approx F(x_m, t)$ is defined at grid points. This also applies to the given velocity with u_m denoting the discrete representation of $u(x_m)$.

A semidiscrete approximation of Equation (2) can be written as

$$\frac{dF_m(t)}{dt} + L_h(F_m(t)) = 0,$$

where L_h is a discrete linear operator associated with the divergence term in each grid point. In this study, both finite difference and finite volume methods commonly used in spectral wave models are discussed to achieve expressions for the discrete operator.

2.2 | A flux-conservative upwind finite difference method

The aim is to present a finite difference scheme which is shock capturing. The divergence term of Equation (2) is approximated by means of a one-sided difference scheme, as follows

$$\left. \frac{\partial uF}{\partial x} \right|_m \approx \frac{u_m F_m - u_{m-1} F_{m-1}}{\Delta x_m} \quad \text{if} \ \ u_m > 0 \tag{3}$$

or

$$\left. \frac{\partial uF}{\partial x} \right|_m \approx \frac{u_{m+1} F_{m+1} - u_m F_m}{\Delta x_{m+1}} \quad \text{if} \ \ u_m < 0. \tag{4}$$

Application of this scheme to the spectral energy balance equation is described in References 2-4.

Remark 1. Strictly speaking, scheme (3)–(4) is neither derived from the local Lax–Friedrichs method nor the approximate Riemann solver of Roe.[8]

For the purpose of the Taylor series analysis of truncation error, we let $h = \max_m \Delta x_m$ be designated as a typical grid size with $\mathcal{O}(h) = \mathcal{O}(M^{-1})$. Based on the reasoning of Turkel,[7] the local truncation error of scheme (3)–(4) for regular grids is then found to be

$$-\frac{1}{2}h\frac{\partial^2 uF}{\partial x^2} + \mathcal{O}(h^2). \tag{5}$$

We can state the following theorem of which the proof is left to the reader.

Theorem 1. *Given a regular grid and regular boundary data, the upwind finite difference scheme (3)–(4) has an optimal rate h of convergence in any norm for IBVP (2). Moreover, the scheme is exact for a constant and a linear flux.*

Remark 2. In the context of spectral wave modeling, it is not uncommon to employ low-order schemes as the curvature of the wave energy flux field usually tends to be small. Additionally, first-order upwind schemes are potentially convenient to counteract the so-called garden sprinkler effect. This effect manifests as nonphysical disintegration of the initial wave field into many separate, smaller wave fields due to the finite resolution of the wave spectrum. Some degree of diffusion (either numerical or artificial) is required to compensate for a lack of dispersion of the continuous spectrum (see, e.g., References 1,9).

Remark 3. Due to the compactness of the stencil of scheme (3)–(4), it can be employed as a locally one-dimensional discrete operator applied to each grid point on an arbitrary mesh in two dimensions, possibly augmented with a local coordinate transformation. Accordingly, Theorem 1 remains true for problems associated to Equation (1). The extension to two dimensions will be dealt with in Section 3.

To demonstrate proper treatment of a shock at the discrete level, we consider a stationary jump in the interval $[x_{m-1}, x_m]$ where the velocity displays a discontinuity over that interval. Furthermore, the velocity is locally positive, that is, $u_m > 0$. Since $duF/dx = 0$ at steady state, flux uF must be continuous across the jump. Note that the preservation of this flux continuity is also reflected in the truncation error (5). Scheme (3) is evidently flux conservative. Indeed, this approximation yields

$$u_m F_m = u_{m-1} F_{m-1}$$

under stationary conditions, implying that flux uF is continuous across the interval $[x_{m-1}, x_m]$. This scheme strictly conserves flux point by point and remains accurate irrespective of the grid size. This is characterized by the following jump condition

$$[uF]_{m-1/2} = u_m F_m - u_{m-1} F_{m-1} = 0 \tag{6}$$

with the operator $[\cdot]_{m-1/2}$ denoting the jump of a flux across $[x_{m-1}, x_m]$. This jump condition is the Rankine–Hugoniot relation expressing the conservation of flux uF. This suggests a zero divergence at point x_m. A similar conclusion can be drawn for the case $u_m < 0$, whereas the velocity is discontinuous across interval $[x_m, x_{m+1}]$.

Remark 4. Since conservation is a necessary ingredient to capture a shock in the solution,[10] the finite difference scheme (3)–(4) is thus conservative.

2.3 | A vertex-centered upwind finite volume scheme

The finite volume method is applied in the following way. It is derived from a weak form of Equation (2) given by

$$\int_\Omega \frac{\partial F}{\partial t} dx + uF\Big|_{\partial\Omega} = 0 \tag{7}$$

with $\partial\Omega$ the boundary of Ω. Since the unknowns are associated with the grid points, a vertex-centered finite volume approach is followed to achieve a conservative discretization. To this end, a dual grid is defined consisting of nonoverlapping control volumes of which their edges lie halfway the grid points. Thus, the control volumes are $\Omega_m = x_{m+1/2} - x_{m-1/2}$,

for $m = 1, \ldots, M-1$, where the volume face is located at $x_{m+1/2} = \frac{1}{2}(x_m + x_{m+1})$. As a consequence, any grid point x_m in the interior of the primal grid represents the center of the control volume. At the boundaries we have $\Omega_0 = x_{1/2} - x_0$ and $\Omega_M = x_M - x_{M-1/2}$. As usual for finite volume schemes, the conserved quantity $F_m(t)$ inside Ω_m is treated as a piecewise constant average.

The weak form (7) is applied to each control volume in the interior of the dual mesh, yielding

$$\Omega_m \frac{dF_m}{dt} = (uF)\Big|_{x_{m-1/2}} - (uF)\Big|_{x_{m+1/2}}, \tag{8}$$

where the volume integral has been approximated by means of the midpoint rule. Both u and F on volume face $x_{m+1/2}$ must be interpolated between nodal values. The approach proposed in References 9,11-13 consists of a first-order upwind discretization for the transported quantity F and a central approximation for the transport velocity u. Thus, we employ

$$u_{m+1/2} \approx \frac{1}{2}(u_m + u_{m+1}), \tag{9}$$

whereas the solution $F_{m+1/2}$ is approximated by means of the following upwind scheme

$$F_{m+1/2} \approx \begin{cases} F_m, & \text{if } u_{m+1/2} > 0 \\ \\ F_{m+1}, & \text{if } u_{m+1/2} < 0 \end{cases}. \tag{10}$$

For a steady state with positive velocity, we end up with the following scheme

$$0 = \frac{(u_{m-1} + u_m)F_{m-1} - (u_m + u_{m+1})F_m}{x_{m+1} - x_{m-1}}. \tag{11}$$

The associated local truncation error is given by

$$-\frac{1}{2}h\left[u\frac{\partial^2 F}{\partial x^2} + \frac{\partial u}{\partial x}\frac{\partial F}{\partial x}\right] + \mathcal{O}(h^2). \tag{12}$$

Although the resulted finite volume scheme is first-order accurate on (quasi-)uniform meshes, it is not exact for either a constant or a linear flux. Furthermore, variations in u and F, particularly across shocks, tend to be far stronger than changes in flux uF (cf. Equation 5). Their smoothness properties can negatively affect errors associated with inadequate resolution of gradients. This is demonstrated in the next two sections.

2.4 | Analysis of the discretization error

We analyze the discretization error of the finite volume scheme for the case with a smooth velocity and the case with a nonsmooth velocity. We treat these cases in $\Omega \subset \mathbb{R}^d$, $d \geq 1$ is the space dimension of the problem. Next, we consider a regular mesh of the domain Ω, either structured or unstructured, with a sufficiently large number of grid vertices. Let this number be $N = \mathcal{O}(M^d)$. The associated dual mesh consists of $\mathcal{O}(N)$ disjoint control volumes Ω_m. The diameter of each volume is given by h_m. Let $h = \max_m h_m$ be the representative grid size so that $|\Omega_m| \leq h^d$ and $\mathcal{O}(h) = \mathcal{O}(M^{-1})$.

To measure the global truncation error, both the discrete L^2-norm and L^∞-norm are used in this study. The L^2-norm is defined as

$$\|\tau\|_2 = \left(\sum_{m=0}^{N} \tau_m^2 |\Omega_m|\right)^{1/2}$$

and the L^∞-norm is

$$\|\tau\|_\infty = \max\{|\tau_0|, \ldots, |\tau_N|\}$$

with $\tau_m = dF(x_m, t)/dt + L_h(F(x_m, t))$ the local truncation error. Furthermore, the discretization error is defined as $e_m = F(x_m, t) - F_m(t)$. Since an upwind approach has been applied, the associated operator L_h is assumed to be stable and consequently uniformly bounded in h. This implies that the estimate $\|e\| \leq k \|\tau\|$ holds with k a positive constant independent of the mesh size and $\|\cdot\|$ a suitable norm.

If both u and F are sufficiently smooth, that is, their derivatives are bounded, then from Equation (12) we have $\tau_m = \mathcal{O}(h)$. The L^∞-norm and L^2-norm of the global truncation error are computed as follows

$$\|\tau\|_\infty = \mathcal{O}(h)$$

and

$$\|\tau\|_2^2 \leq h^d \sum_m \tau_m^2 \sim \mathcal{O}(h^d)\,\mathcal{O}(M^d)\,\mathcal{O}(h^2) = \mathcal{O}(h^2).$$

Hence, $\|\tau\|_2 = \mathcal{O}(h)$. The global order of accuracy is 1, and so is the discretization error. In fact, this holds in any norm.

Next, we continue to analyze the case when the discrete function u_m exhibits a discontinuity across the interval $[x_{m-1}, x_m]$, that is, $u_m - u_{m-1}$ is away from zero while $h \to 0$. Since u_m and u_{m-1} are bounded on this interval, its first derivative tends to behave like $\mathcal{O}(h^{-1})$. This implies that truncation error (12) is increased to $\tau_m = \mathcal{O}(1)$. We assume that the number of intervals over which the velocity is not smooth is $\mathcal{O}(M^{d-1})$. Consequently, the global order of accuracy in L^∞ and L^2 are, respectively,

$$\|e\|_\infty \sim \|\tau\|_\infty = \mathcal{O}(1)$$

and

$$\|\tau\|_2^2 \leq h^d \sum_m \tau_m^2 \sim \mathcal{O}(h^d)[\mathcal{O}(M^d)\,\mathcal{O}(h^2) + \mathcal{O}(M^{d-1})\,\mathcal{O}(1)] = \mathcal{O}(h^d)\,\mathcal{O}(M^{d-1}) = \mathcal{O}(h)$$

so that $\|e\|_2 \sim \|\tau\|_2 = \mathcal{O}(h^{1/2})$, that is, the global order of accuracy is 0.5. Hence, the following theorem has been obtained.

Theorem 2. *Given a regular grid and regular boundary data, the global order of accuracy of the vertex-centered upwind finite volume scheme (8)–(10) applied to IBVP (2) is 1 if the velocity is smooth, but reduces to 0.5 in L^2-norm and 0 in L^∞-norm when the velocity is a nonsmooth function.*

Remark 5. The apparent suboptimal convergence rate of $h^{1/2}$ in the L^2-norm caused by the lack of smoothness of the velocity field is substantiated by the work of Bouche et al.,[14] who reported that the loss of accuracy of the upwind finite volume method applied to a linear scalar advection equation on arbitrary meshes is due to the irregularity of the solution. Indeed, since in our case the velocity field is nonsmooth, so is the solution. Note that the study described in Reference 14 assumes a Lipschitz continuous velocity field, while the smoothness properties of solutions solely depend on those of initial and boundary data involved.

Remark 6. Despite the fact that finite volume scheme (8) using Equations (9) and (10) is (locally and globally) conservative by construction (on dual grid), it does not converge to a solution of the desired accuracy, provided a given nonsmooth u. The underlying rationale for this anomaly is that the discontinuity in the velocity is not captured by the dual mesh in the optimal sense. Observe that this finding does not conflict with the Lax–Wendroff theorem,[15] that is, the associated solution converges to a weak solution of Equation (2) as summation by parts can still be applied.

2.5 | Analysis of the conservation error

From scheme (11) we derive the following jump condition

$$[uF]_{m-1/2} = u_m\,F_m - u_{m-1}\,F_{m-1} = u_m\,F_{m-1} - u_{m+1}\,F_m. \tag{13}$$

Since the velocity field is not divergence free, the right-hand side of Equation (13) gives rise to internal sources of F_m. This implies $u_m F_m \neq u_{m-1} F_{m-1}$, that is, continuity of flux uF across the primal grid cell is lost. To appreciate this observation, it should be pointed out that the magnitude of the error due to disbalance (13) can be significant, especially when a nonsmooth velocity is dealt with.

A local conservation error is the deviation of scheme (8)–(10) on primal mesh from a semidiscrete form of Equation (2) satisfying the Rankine–Hugoniot jump condition (6). Let us consider the case with positive velocities everywhere. Furthermore, it suffices to choose a uniform grid with constant grid size Δx. We then examine the following semidiscrete scheme (cf. Equation 11)

$$\frac{dF_m}{dt} + \frac{(u_m + u_{m+1})F_m - (u_{m-1} + u_m)F_{m-1}}{2\Delta x} = 0.$$

This equation can be rewritten as

$$\frac{dF_m}{dt} + \frac{u_m F_m - u_{m-1} F_{m-1}}{\Delta x} = \frac{(u_m - u_{m+1})F_m - (u_{m-1} - u_m)F_{m-1}}{2\Delta x}. \tag{14}$$

As noted earlier, the solution of scheme (8)–(10) does not necessarily approach the sought solution of Equation (2) upon convergence. Instead, it is a solution of the following modified equation

$$\frac{\partial F}{\partial t} + \frac{\partial uF}{\partial x} = \Sigma \tag{15}$$

with Σ a nonconservative source term. Assuming that scheme (14) is consistent with Equation (15) and making use of Taylor expansions, we obtain

$$\Sigma = -\frac{1}{2}h\left(F\frac{\partial^2 u}{\partial x^2} + \frac{\partial u}{\partial x}\frac{\partial F}{\partial x}\right)$$

so that a nonsmooth u yields a local conservation error of $\mathcal{O}(1)$. However, if the velocity is sufficiently smooth, the conservation error vanishes as the mesh size tends to zero. Note that Σ has a compact support.

Remark 7. It is of crucial importance to observe that the situation becomes totally different for the case of zero divergence velocity field, that is, $du/dx = 0$. This implies that F must be a constant because $dF/dx = 0$. Since F_{m-1} and F_m are involved in Equation (11) and because of consistency, we have $F_m = F_{m-1}$. Consequently, the right-hand side of Equation (13) is zero. In addition, according to Equation (12), smoothness of u is not relevant in this regard. Note that Equation (12) is then identical to Equation (5) and $\Sigma \equiv 0$, implying that scheme (8)–(10) is suitable for the case of divergence-free velocity fields.

3 | DISCRETIZATIONS IN TWO SPACE DIMENSIONS

A two-dimensional bounded domain $\Omega \subset \mathbb{R}^2$ is discretized using a partitioned mesh that is either structured (quadrilaterals) or unstructured (triangles). The discussed methods of Section 2 can be readily generalized to these types of meshes. In this article, we restrict ourselves to triangular grids. We consider a triangulation of Ω as depicted in Figure 1. The discrete solution and the discrete velocity are positioned at the grid vertex and Equation (1) is solved in each vertex. The next two sections reconsider the upwind finite difference method and the vertex-centered upwind finite volume schemes in two dimensions, respectively, to approximate the divergence term of Equation (1) to be derived shortly.

3.1 | The flux-conservative upwind finite difference scheme

We seek for the solution at vertex "c" in Figure 1. The spatial discretization involves each cell of this vertex. Let us consider a triangle $\triangle 123$ as shown in Figure 2. The discrete solution at vertices 1, 2, and 3 are denoted by F_1, F_2, and F_3, respectively. The discretization comprises a coordinate transformation $\boldsymbol{x}(\boldsymbol{\xi})$ with $\boldsymbol{\xi} = (\xi, \eta)$ a local coordinate system and $\boldsymbol{x} = (x, y)$ the
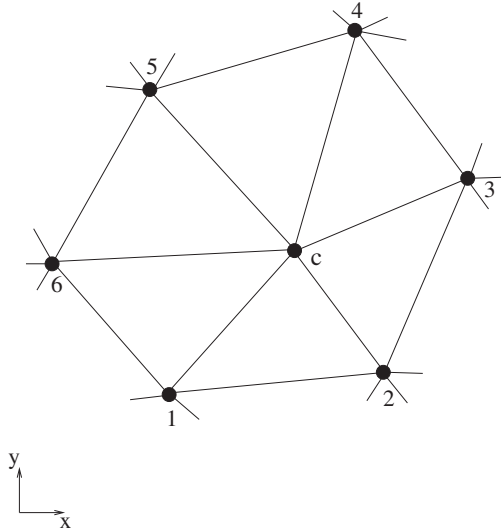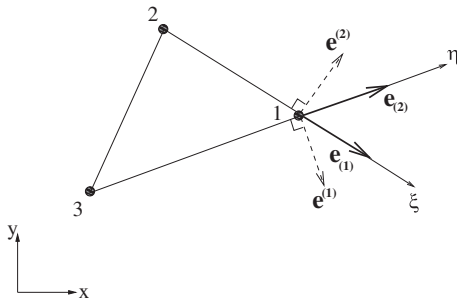
FIGURE 1 An example of triangulation

FIGURE 2 A triangular cell with geometrical quantities used for spatial discretization. Definitions of these quantities are provided in the text

Cartesian one. The base vectors $\boldsymbol{e}_{(1)}$ and $\boldsymbol{e}_{(2)}$ are tangential to the coordinate lines ξ and η, and are given by

$$\boldsymbol{e}_{(1)} = \frac{\partial \boldsymbol{x}}{\partial \xi}, \quad \boldsymbol{e}_{(2)} = \frac{\partial \boldsymbol{x}}{\partial \eta}.$$

The mapping $\boldsymbol{x}(\boldsymbol{\xi})$ is piecewise bilinear. Furthermore, we choose the mapping such that $\Delta\xi = \Delta\eta = 1$. Referred to Figure 2, the base vectors can be computed as

$$\boldsymbol{e}_{(1)} = \boldsymbol{x}_1 - \boldsymbol{x}_2, \quad \boldsymbol{e}_{(2)} = \boldsymbol{x}_1 - \boldsymbol{x}_3$$

with $\boldsymbol{x}_i = (x_i, y_i)$ the position vector of vertex i. The vectors

$$\boldsymbol{e}^{(1)} = \nabla\xi, \quad \boldsymbol{e}^{(2)} = \nabla\eta$$

are normal to the coordinate line of constant ξ and η, respectively, and are found to be

$$\boldsymbol{e}^{(1)} = \frac{1}{D}(e_{(2)}^2, -e_{(2)}^1)^\top, \quad \boldsymbol{e}^{(2)} = \frac{1}{D}(-e_{(1)}^2, e_{(1)}^1)^\top, \quad D = e_{(2)}^2 e_{(1)}^1 - e_{(1)}^2 e_{(2)}^1.$$

We proceed with the discretization. First, we expand the divergence term of Equation (1)

$$\nabla \cdot (\boldsymbol{v}\, F) = \frac{\partial uF}{\partial x} + \frac{\partial vF}{\partial y}.$$

Using the chain rule, we obtain

$$\nabla \cdot (\boldsymbol{v}F) = e_1^{(1)} \frac{\partial uF}{\partial \xi} + e_1^{(2)} \frac{\partial uF}{\partial \eta} + e_2^{(1)} \frac{\partial vF}{\partial \xi} + e_2^{(2)} \frac{\partial vF}{\partial \eta}.$$

Next, we approximate the derivatives by means of a one-sided first-order difference scheme

$$\frac{\partial uF}{\partial \xi} \approx (uF)_1 - (uF)_2, \quad \frac{\partial uF}{\partial \eta} \approx (uF)_1 - (uF)_3$$

and

$$\frac{\partial vF}{\partial \xi} \approx (vF)_1 - (vF)_2, \quad \frac{\partial vF}{\partial \eta} \approx (vF)_1 - (vF)_3,$$

if velocity vector $\boldsymbol{v}$ is contained in triangle $\triangle 123$ of Figure 2, that is, $\boldsymbol{v} \cdot \boldsymbol{e}^{(1)} > 0$ and $\boldsymbol{v} \cdot \boldsymbol{e}^{(2)} > 0$, otherwise the grid cell under consideration is skipped. The approximation is completed after substitution

$$\nabla \cdot (\boldsymbol{v}F) \approx uF\big|_2^1 e_1^{(1)} + uF\big|_3^1 e_1^{(2)} + vF\big|_2^1 e_2^{(1)} + vF\big|_3^1 e_2^{(2)}. \tag{16}$$

This spatial discretization is first-order accurate, has a compact stencil, and conserves $\boldsymbol{v}\,F$ locally.

The latter is proved as follows. We integrate the divergence term on the left-hand side of Equation (16) over triangle $\triangle 123$ and subsequently consider the following surface integral

$$\oint_{\partial\triangle} \boldsymbol{v}\,F \cdot d\boldsymbol{s} \approx \sum_e (\boldsymbol{v}\,F)_e \cdot d\boldsymbol{s}_e$$

with the summation running over the three edges of the triangle and $d\boldsymbol{s}_e$ is the edge vector pointing outward of which the length equals the corresponding edge. Note that the sum of the edge vectors is zero. The vectors $\boldsymbol{v}\,F$ at the midpoints of the edges are taken as arithmetic means, which are exact owing to the bilinear mapping. Referred to Figure 2, the summation over the edges yields

$$-\frac{1}{2}[(\boldsymbol{v}\,F)_1 + (\boldsymbol{v}\,F)_2] \cdot \boldsymbol{e}^{(2)} + \frac{1}{2}[(\boldsymbol{v}\,F)_2 + (\boldsymbol{v}\,F)_3] \cdot (\boldsymbol{e}^{(1)} + \boldsymbol{e}^{(2)}) - \frac{1}{2}[(\boldsymbol{v}\,F)_3 + (\boldsymbol{v}\,F)_1] \cdot \boldsymbol{e}^{(1)}$$

$$= -\frac{1}{2}[(\boldsymbol{v}\,F)_1 \cdot \boldsymbol{e}^{(1)} + (\boldsymbol{v}\,F)_1 \cdot \boldsymbol{e}^{(2)} - (\boldsymbol{v}\,F)_2 \cdot \boldsymbol{e}^{(1)} - (\boldsymbol{v}\,F)_3 \cdot \boldsymbol{e}^{(2)}] = 0.$$

The equality to zero is by virtue of requiring a divergence-free flux $\boldsymbol{v}\,F$ and discretization (16), meaning that it is source free.

Remark 8. Flux $\boldsymbol{v}\,F$ is continuous along the characteristic contained in triangle $\triangle 123$.

Remark 9. The present discretization is akin to the streamline upwind finite element technique (see, e.g., Reference 16) and is also implemented in a spectral wave model as reported in Reference 17.

3.2 | The vertex-centered upwind finite volume scheme

The chosen control volume around each vertex is the centroid dual as depicted in Figure 3. It is constructed by joining the centroids neighboring the vertex under consideration. The set of dual volumes must fill the entire domain and must also be nonoverlapping. These volumes form the dual mesh. In the following we use the notation from Figure 3.

Integrating Equation (1) over a control volume yields

$$\int_{\Omega_c} \frac{\partial F}{\partial t} d\Omega + \oint_{\partial\Omega_c} (\boldsymbol{v}\,F \cdot \boldsymbol{n}) ds = \int_{\Omega_c} Sd\Omega,$$

where Ω_c is the area of the control volume and $\boldsymbol{v}\,F \cdot \boldsymbol{n}$ is the normal flux at the edge of the control volume. The surface integral is approximated as

$$\oint_{\partial\Omega_c} (\boldsymbol{v}\,F \cdot \boldsymbol{n}) ds \approx \sum_j \boldsymbol{v}_j \cdot \boldsymbol{n}_j \, F_j ds_j$$
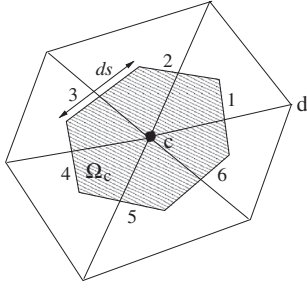
FIGURE 3 Control volume (centroid dual) of vertex is shaded. Some notation is introduced and clarified in the text

with the summation over each edge j of the control volume and ds_j is the length of the edge. The semidiscretized equation reads

$$\Omega_c \frac{\partial F}{\partial t} + \sum_j \boldsymbol{v}_j \cdot \boldsymbol{n}_j \, F_j ds_j = \Omega_c S,$$

where the volume integrals have been approximated using the midpoint rule.

As outlined in Section 2.3, the velocity component normal to the volume edge is obtained from arithmetic averaging. Referring to Figure 3, we have, for instance,

$$(\boldsymbol{v} \cdot \boldsymbol{n})_1 = \frac{1}{2}(\boldsymbol{v}_c + \boldsymbol{v}_d) \cdot \boldsymbol{n}_1, \tag{17}$$

whereas the unknown F_1 equals the value at vertex that is located upwind of the volume edge. Hence

$$F_1 \approx \begin{cases} F_c, & \text{if } (\boldsymbol{v} \cdot \boldsymbol{n})_1 > 0 \\ F_d, & \text{if } (\boldsymbol{v} \cdot \boldsymbol{n})_1 < 0 \end{cases}. \tag{18}$$

4 | NUMERICAL EXPERIMENTS

This section presents results of the grid convergence tests for the following semidiscrete schemes

1. **FDM-flux** The flux-conservative finite difference scheme (16).
2. **FVM-upw/cent** The vertex-centered upwind finite volume scheme (18) with the central velocity approximation (17).

Both schemes have been tested on the following equation

$$\frac{\partial F}{\partial t} + \frac{\partial uF}{\partial x} + \frac{\partial vF}{\partial y} = 0.$$

The computational domain was $\Omega = \{(x, y) | 0 \leq x \leq 3000, \ 0 \leq y \leq 15,000\}$. The initial and boundary conditions were given by $F(x, y, 0) = 0$ and $F(0, y, t) = 1$, respectively, $(x, y) \in \Omega$, $t > 0$. The velocity field was given by $(u, v) = (v \cos \theta, v \sin \theta)$ with θ the direction of the velocity vector measured counterclockwise with respect to the x-axis. We have considered two cases. In the first test problem, the velocity field was given as a continuously differentiable function

$$v(x) = \left(1550 - \frac{1}{2}x\right)^{1/3}.$$

The corresponding steady-state solution reads

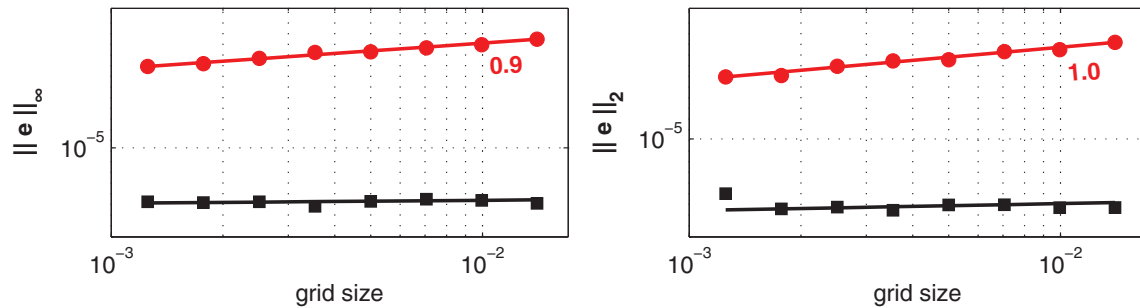$$F(x) = \frac{\sqrt[3]{1550}}{v(x)}.$$

FIGURE 4 Comparison of spatial accuracy and convergence rate for the smooth velocity case obtained with different schemes: **FDM-flux** (black squares) and **FVM-upw/cent** (red dots). The numbers in the plots indicate the rate of convergence [Colour figure can be viewed at wileyonlinelibrary.com]

The second test case is dealt with a nonsmooth velocity field, as follows

$$v(x, y) = \begin{cases} 12, & 30\,x - y < 37{,}500 \\ 4, & 30\,x - y > 37{,}500 \end{cases}$$

and the associated steady-state solution is

$$F(x, y) = \frac{12}{v(x, y)}.$$

A series of eight uniform triangular meshes has been selected, with the number of nodes increasing from 5100 to 640,000 through doubling. The semidiscrete schemes were integrated in time using a fourth-order explicit Runge–Kutta four-stage method. A sufficiently small time step was chosen such that for all grids the temporal error is much smaller than the error related to the spatial discretization. In addition, the fully discrete equations were found to be stable. All the simulations were terminated when the steady-state solution was obtained. The spatial accuracy and convergence rate for the two schemes were evaluated using the discrete L^∞ and L^2 error norms. The errors were determined by comparing the numerical solutions to the exact solution. Finally, in anticipation of the findings, the measured errors appeared to be insensitive to the velocity direction θ. For this reason, the results presented below correspond to only one arbitrary $\theta \in (-\frac{1}{2}\pi, \frac{1}{2}\pi)$.

Figure 4 illustrates the behavior of the global errors as function of the grid size for the first test case with the smooth velocity. A linear convergence for scheme **FVM-upw/cent** was observed as theoretically predicted by Theorem 2. The figure also reveals that increasing resolution did not lead to accuracy improvement for scheme **FDM-flux** as it is already exact up to machine precision (in line with Theorem 1). Moreover, this scheme outperformed **FVM-upw/cent** with an average difference in error being about five orders of magnitude (10^{-7} vs. 10^{-2}).

The results of the case with the nonsmooth velocity are displayed in Figure 5. Like the previous case, for each individual grid scheme **FDM-flux** exhibited a global error of machine precision. On the other hand, there is no convergence in L^∞ for **FVM-upw/cent**, whereas in L^2 convergence drops towards 0.6.

The discussed experiments confirmed the results from the analysis outlined in Section 2.4. The next section demonstrates some consequences of the present findings for an application of practical interest.

5 | APPLICATION

This section deals with a synthetic test case concerning wave shoaling and wave refraction in the presence of variable bathymetry. In particular, these wave processes are due to submerged shoals. Shoals are a topographic feature of the coastal waters where wind waves and swells are influenced by the relatively shallow water depths. Examples are the submerged ridges, banks, and bars. On approaching a shoal, waves are subject to different processes that are mainly determined by the bed variations. An increase in the wave energy around the shoal occurs due to shoaling when waves experience the
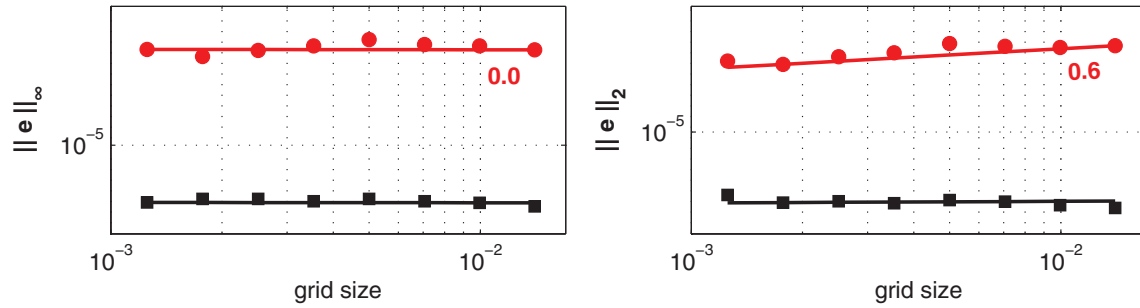
FIGURE 5 Comparison of spatial accuracy and convergence rate for the nonsmooth velocity case obtained with different schemes: **FDM-flux** (black squares) and **FVM-upw/cent** (red dots). The numbers in the plots indicate the rate of convergence [Colour figure can be viewed at wileyonlinelibrary.com]

decreasing depth, and due to refraction induced by the local changes in the depth along the wave crests.[1] These effects upon waves can be significant when there are abrupt changes in water depths over a small distance.

The governing equation reads

$$\nabla \cdot (\boldsymbol{c}_g E) = -\frac{\partial c_\theta E}{\partial \theta}, \tag{19}$$

where θ is the wave direction (measured counterclockwise from the positive x-axis in radians), $E(\boldsymbol{x}, \theta)$ is the steady-state wave energy (in m²/rad), $\boldsymbol{c}_g = (c_g \cos \theta, c_g \sin \theta)$ is the propagation velocity of a wave group along a wave ray, and c_θ is the rate of change of wave direction (in rad/s). The magnitude of the group velocity $c_g \equiv \partial \sigma / \partial k$ can be calculated using the following dispersion relation of the linear wave theory[1]

$$\sigma^2 = gk \tanh(kd)$$

with $\sigma = 2\pi f$ the radian (intrinsic) frequency, k the wave number, g the gravitational acceleration and $d(\boldsymbol{x})$ the water depth.

Equation (19) represents the local balance between wave propagation, shoaling, and refraction under stationary conditions. This equation is conceptually ideal to study the evolution of swell waves in isolation from the complex wave dynamics due to local generation by wind, nonlinear interactions and wave dissipation. The term on the left-hand side of Equation (19) is the divergence of the wave energy flux per unit crest length $\boldsymbol{c}_g E$. This term can be rewritten as

$$\boldsymbol{c}_g \cdot \nabla E + (\nabla \cdot \boldsymbol{c}_g) E$$

with the first term expressing the propagation of wave energy with the group velocity along a wave ray, and the second term describing the effect of wave shoaling. The term on the right-hand side of Equation (19) is the refraction term specifying the turning of waves due to local depth changes. In particular, the depth dependence is incorporated in the turning rate c_θ. Details on the computation of the refraction term can be found in Reference 4. It suffices to mention that this term is approximated by means of a first-order flux-conservative upwind finite difference method and remains physically accurate under the condition of mild and moderate bottom slopes.

The objective of the present test case is to assess the accuracy and conservation properties of the two discussed methods for the solution of Equation (19), **FDM-flux** and **FVM-upw/cent**. In particular, we conjecture that a local conservation error that arise in the approximation of the divergence term of Equation (19), due to scheme **FVM-upw/cent**, would create an imbalance between shoaling and refraction. This implies a different spatial distribution of the wave energy. The semidiscrete schemes were artificially marched in time using the fourth-order Runge–Kutta method until a steady-state solution was reached.

The schematization of the test case is taken from Reference 18, though the physical and environmental conditions have been adapted to emphasize the finite depth effects. It consists of an area of 10 km × 10 km with two submerged, crescent-shaped shoals, the largest spanning 2 km and the smallest about 1 km. The topography of the considered shoals
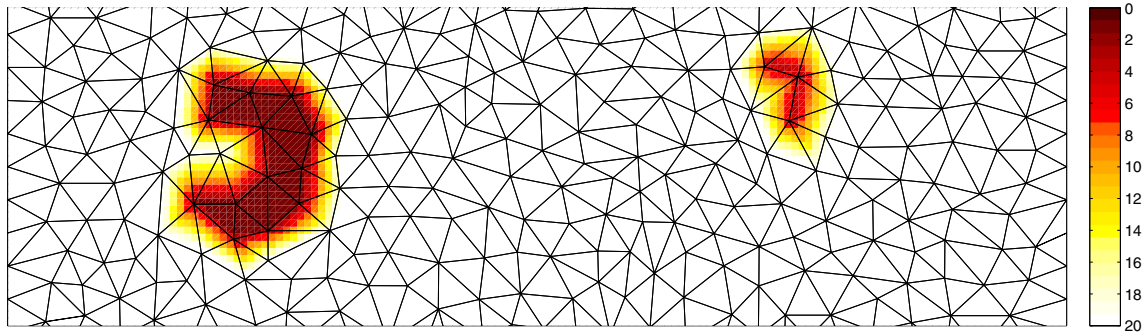
FIGURE 6 A part of the model domain with the bathymetry (in m) projected on the triangular mesh, displaying two submerged shoals [Colour figure can be viewed at wileyonlinelibrary.com]

is depicted in Figure 6. The surrounded bed is 20 m deep. The summit of the shoals is relatively shallow: 1.5 and 3.5 m deep at the center of the large and small one, respectively.

The unstructured grid employed in the test case consists of 827 nodes and 1504 triangles. The grid size varies in between 100 and 400 m. This mesh resolution should provide an economical representation of the bathymetric variability in this region, while a numerical scheme must remain accurate in the presence of strong depth gradients.

A monochromatic, long-crested swell wave with a period of 15 s enters the area at the southern boundary of the computational domain with a wave direction pointing northward, and allows to propagate through to the northern boundary. The amount of incident wave energy is constant along the boundary and equals $E = 1 \text{ m}^2$.

First, we consider the case of only wave shoaling, that is, the right-hand side of Equation (19) is identically zero. Consequently, the wave direction remains constant. In this respect, the shoaling coefficient is the intended quantity to predict. It is defined as $K_s(\boldsymbol{x}) = \sqrt{c_{g,\infty}/c_g(\boldsymbol{x})}$ with $c_{g,\infty}$ the group velocity at deep water.[1] Owing to flux conservation, the shoaling coefficient can be verified using $K_s(\boldsymbol{x}) = \sqrt{E(\boldsymbol{x})}$. (Note that wave energy at deep water is 1 m^2.) The group velocity at deep water is 11.7 m/s according to the dispersion relation. The group velocity on top of the large and small shoals is 3.8 and 5.7 m/s, respectively. Thus, the shoaling coefficient corresponding to the center of the large shoal is 1.75 and to the center of the small one is 1.43. As a consequence, the amount of wave energy on top of the shallower shoal is larger than on top of the deeper shoal.

Figures 7 and 8 show the results of both methods. Scheme **FDM-flux** displays correctly the spatial distribution of the shoaling coefficient across the shoals as it changes *locally* with water depth (cf. Figure 6). Also, it predicts a shoaling coefficient of 1.75 at the large shoal and 1.43 at the small one. In contrast, scheme **FVM-upw/cent** underestimates the amount of wave energy fairly, whereas its spatial distribution is incorrect. The predicted shoaling coefficients on top of the large and small shoals appeared to be 1.54 and 1.1, respectively. This example demonstrates that this vertex-centered finite volume scheme, applied to the divergence term of Equation (19), does not necessarily contribute to correct wave
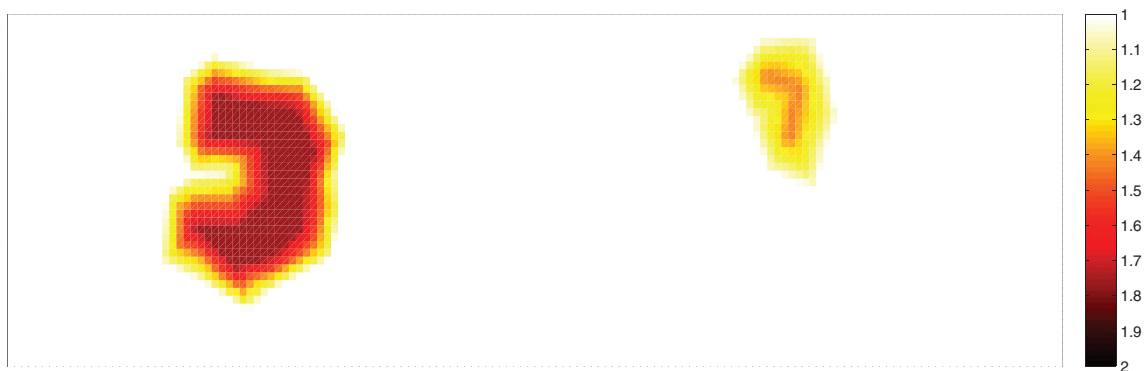

FIGURE 7 Spatial distribution of shoaling coefficient obtained with scheme **FDM-flux** [Colour figure can be viewed at wileyonlinelibrary.com]
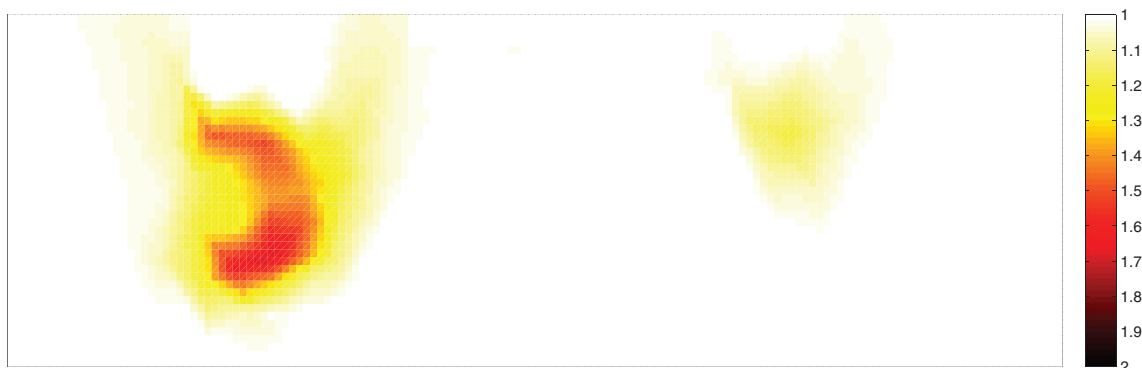
FIGURE 8 Spatial distribution of shoaling coefficient obtained with scheme **FVM-upw/cent** [Colour figure can be viewed at wileyonlinelibrary.com]

shoaling over steep bottom gradients. This is reflected by the zeroth-order convergence rate in L^∞. Thus we may conclude that the scheme is practically inconsistent, even though the convergence rate in L^2 is larger than 0.

The next simulations are the ones with depth refraction included. Refraction leads to a change of the wave direction across the shoals and its effect is stronger in shallower water than in deeper water. Consequently, the swell wave turns towards the shoal. In addition, refraction causes the directional width of the wave group to change along the curved trajectory.[11]

In the context of interpreting the simulation results, the wave energy in each location is obtained by integration over all spectral directions

$$E(\boldsymbol{x}) = \int_0^{2\pi} E(\boldsymbol{x}, \theta) d\theta$$

and the wave direction is the mean of all the individual wave directions occurred in the spectrum and is computed as[5]

$$\overline{\theta}(\boldsymbol{x}) = \arctan\left(\frac{\int_0^{2\pi} \sin\theta \; E(\boldsymbol{x}, \theta) d\theta}{\int_0^{2\pi} \cos\theta \; E(\boldsymbol{x}, \theta) d\theta}\right).$$

Figure 9 displays the results of **FDM-flux** and demonstrates clearly the effect of wave turning at shallower depths. In addition, there is a convergence of wave energy when the swell approaches the shoal and a divergence of wave energy when it leaves the shoal. This is consistent with the Snel's law.[1] As a result, the distance between the two neighboring rays decreases on top of the shoal, suggesting an increase in wave energy on-site. Furthermore, the effect of refraction at the deeper shoal is relatively much weaker. These findings are in line with the geometric optics.[1,5]
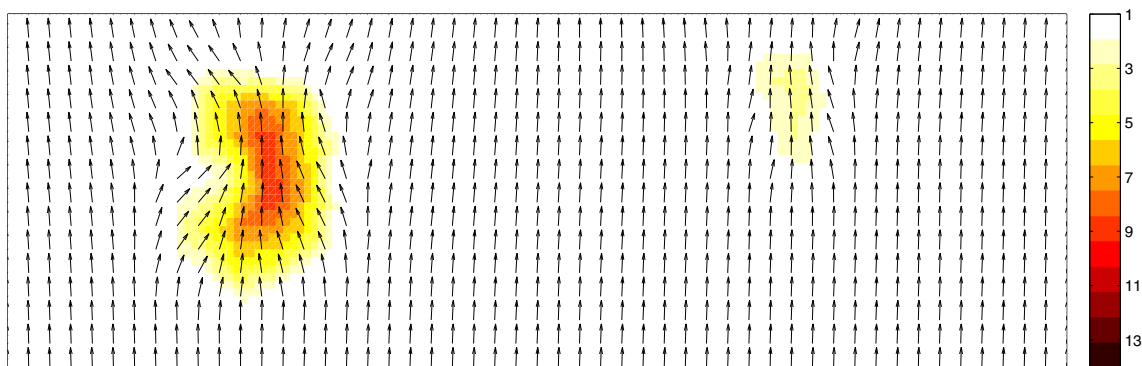

FIGURE 9 Spatial distribution of wave energy (in m^2) and vector plot of mean wave direction obtained with scheme **FDM-flux** [Colour figure can be viewed at wileyonlinelibrary.com]
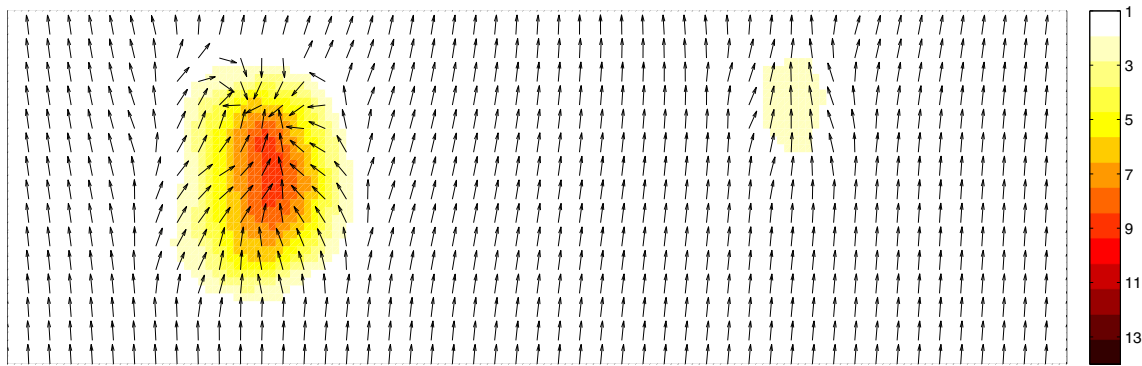
FIGURE 10 Spatial distribution of wave energy (in m²) and vector plot of mean wave direction obtained with scheme **FVM-upw/cent** [Colour figure can be viewed at wileyonlinelibrary.com]

The results achieved with scheme **FVM-upw/cent** are depicted in Figure 10. An unnatural pattern is present where the wave turning is steadily magnified at the large shoal. It is striking that the swell wave even propagates backward towards the rear side of the large shoal. There is also a more symmetric refraction pattern found around both irregular shoals compared to the other method. These observed deviations can be attributed to the lack of flux conservation of the **FVM-upw/cent** method, generating a local conservation error which evidently disbalance the interplay between shoaling and refraction, subsequently leading to an inaccurate spatial distribution of wave energy and hence refraction pattern. This conclusion is implied by the analysis outlined in Section 2.5.

6 | CONCLUSIONS

An accuracy study of two semidiscrete schemes to approximate the divergence term of a transport equation with nonzero divergence velocity field has been carried out. These schemes were originally introduced in the field of spectral modeling of wind waves and swells in coastal seas. They differ in the manner of computing the scalar flux. A convergence analysis of these methods has not been discussed in the literature. The present article has aimed to contribute to such an analysis.

The first scheme is an upwind finite difference scheme which proved to be flux conservative and converges with a rate of $\mathcal{O}(h)$ irrespective of the smoothness of the velocity. It has been further demonstrated that this scheme is exact when the flux is constant or linear in the $\boldsymbol{x}$-space.

The second scheme is a vertex-centered upwind finite volume scheme with central discretization of the velocity. This scheme has shown to fail to conserve the flux locally and has also an adverse effect on the discretization error. The scheme proved to be first-order convergent in any norm for smooth velocity field, but this order of convergence drops to 0.5 in L^2-norm and 0 in L^∞-norm when considering nonsmooth velocity.

Numerical experiments have been conducted that have confirmed the abovementioned findings. Finally, a wave model test with irregular topography has been undertaken which shows that a lack of flux conservation has a significant impact on wave shoaling and wave refraction, especially at shallower depths.

ORCID

Marcel Zijlema https://orcid.org/0000-0001-9020-2105

REFERENCES

1. Holthuijsen LH. *Waves in Oceanic and Coastal Waters*. Cambrideg: Cambridge University Press; 2007.
2. The WAMDI Group. The WAM model - a third generation ocean wave prediction model. *J Phys Oceanogr*. 1988;18:1775-1810.
3. Tolman HL. A third-generation model for wind waves on slowly varying, unsteady and inhomogeneous depths and currents. *J Phys Oceanogr*. 1991;21:782-797.
4. Booij N, Ris RC, Holthuijsen LH. A third-generation wave model for coastal regions 1. model description and validation. *J Geophys Res*. 1999;104(C4):7649-7666.
5. Komen GJ, Cavaleri L, Donelan M, Hasselmann K, Hasselmann S, Janssen PAEM. *Dynamics and Modelling of Ocean Waves*. Cambridge: Cambridge University Press; 1994.

6. Boyer F. Analysis of the upwind finite volume method for general initial- and boundary-value transport problems. *IMA J Numer Anal.* 2012;32:1404-1439.

7. Turkel E. Accuracy of schemes with nonuniform meshes for compressible fluid flows. *Appl Numer Math.* 1986;2:529-550.

8. LeVeque RJ. *Finite Volume Methods for Hyperbolic Problems.* Cambridge: Cambridge University Press; 2004.

9. Janssen PAEM. Progress in ocean wave forecasting. *J Comput Phys.* 2008;227:3572-3594.

10. Hou TY, Le Floch PG. Why nonconservative schemes converge to wrong solutions: error analysis. *Math Comput.* 1994;62:497-530.

11. Hubbert KP, Wolf J. Numerical investigation of depth and current refraction of waves. *J Geophys Res.* 1991;96(C2):2737-2748.

12. Christensen ED, Johnson M, Sørensen OR, Hasager CB, Badger M, Larsen SE. Transmission of wave energy through an offshore wind turbine farm. *Coast Eng.* 2013;82:25-46.

13. The WAVEWATCH III® Development Group. *User Manual and System Documentation of WAVEWATCH III® Version 6.07 Technical Note 333.* College Park, MD: NOAA/NWS/NCEP/MMA; 2019.

14. Bouche D, Ghidaglia JM, Pascal F. Error estimate and the geometric corrector for the upwind finite volume method applied to the linear advection equation. *SIAM J Numer Anal.* 2005;43:578-603.

15. Lax PD, Wendroff B. Systems of conservation laws. *Commun Pure Appl Math.* 1960;13:217-237.

16. Shemirani F, Jambunathan K. Conservative monotone streamline upwind formulation using simplex elements. *Int J Numer Methods Fluids.* 1992;14:1245-1257.

17. Zijlema M. Computation of wind-wave spectra in coastal waters with SWAN on unstructured grids. *Coast Eng.* 2010;57:267-277.

18. Huchet M, Leckler F, Filipot JF, et al. On the high resolution coastal applications with WAVEWATCH III®. Paper presented at: Proceedings of the 14th International Workshop on Wave Hindcasting and Forecasting; 2015:1-13; Key West, FL.