Journal of Computational Physics **183,** 142–164 (2002)
doi:10.1006/jcph.2002.7180

Toward an Oscillation-Free, Mass Conservative, Eulerian–Lagrangian Transport Model

Anabela Oliveira[1] and André B. Fortunato

Departamento de Hidráulica, Núcleo de Estuários, Laboratório Nacional de Engenharia Civil, Av. do Brasil 101, 1700-066 Lisboa, Portugal
E-mail: aoliveira@lnec.pt and afortunato@lnec.pt

Received May 2, 2001; revised January 8, 2002

Ten numerical schemes to reduce spurious oscillations in transport simulations were implemented and tested in a Eulerian–Lagrangian control-volume finite element model. The schemes included both new and existing flux-corrected transport (FCT) algorithms and nonlinear filters. The ability of the methods to eliminate numerical oscillations while preserving mass and minimizing the introduction of numerical diffusion was compared in 2D tests of varying complexity. The application of local mass correction algorithms was shown to be vital to avoid the introduction of mass errors by FCT. While none of the methods emerged as optimal for all cases, a new FCT method with a local mass correction scheme and a nonlinear filter can be recommended as the best approaches in general. The method of choice depends on the problem being solved, in particular on the concentration gradients and on the grid resolution.

Key Words: flux-corrected transport methods; nonlinear filters; transport equation.

1. INTRODUCTION

The numerical solution of advective-dominated transport has long faced the dilemma between spurious oscillations and numerical damping. Typically, high-order methods target the preservation of peak concentrations and sharp gradients but suffer from numerical oscillations. In contrast, low-order methods (e.g., first-order upwind methods) offer positivity at the expense of numerical damping.

Clearly, many problems exist in which both positivity and peak preservation are required. For instance, numerical oscillations in a water-quality transport model can lead to negative concentrations, hence to unrealistic behaviors. Simultaneously, numerical damping can be

[1] Fax: +351-21-8443016.

0021-9991/02 $35.00

a serious flaw if there are threshold concentration limits that should not be reached (e.g., due to environmental regulations).

While oscillations indicate the need for additional refinement [1], this refinement may be prevented by the associated computational costs. Adaptive methods (e.g., [2–4]) are a possible alternative. However, their application in complex multidimensional systems may be hampered by the increase in computational costs and the need for consistent interpolation of variables between grids.

Several solutions have been proposed to mitigate or eliminate spurious oscillations with minimal introduction of numerical damping. Among these, flux-corrected transport methods (FCT) [5], the localized use of a low-order method [6], and the use of nonlinear filters [7] focus on the local elimination of the oscillations, to preserve the overall accuracy. A comparison of nonlinear filters and the localized use of a low-order method showed that the former are superior to the latter in both oscillation removal and mass conservation [8].

This paper presents a detailed comparison of nonlinear filters and FCT methods in two-dimensional unstructured grids. Within the FCT concept, we compare previously proposed formulations [5, 9] and propose new formulations that improve both mass conservation and accuracy. A new formulation that combines FCT with nonlinear filter concepts is also proposed and analyzed.

The analysis is conducted in the framework of finite-element-based Eulerian–Lagrangian methods [10]. These methods are very attractive for advection-dominated transport modeling, since they allow Courant numbers larger than 1 while retaining the convenience of a fixed computational grid. Typically, advection is solved by the backward method of characteristics, while diffusion is solved by finite elements, finite differences, or finite volumes. In the past decade, these methods have been substantially improved to enhance mass conservation [11–13], which was their major limitation.

The filter and FCT formulations were implemented in a control-volume finite-element-based Eulerian–Lagrangian model (VELA [14]). This model, which combines some of the most advanced approaches for transport modeling using ELMs, compared favorably with several finite element models [15], ranking therefore among the most accurate transport formulations available. For the purpose of this work, though, the choice of the specific model used to implement the oscillation-removal schemes is not very relevant. This model was selected for its high accuracy and, therefore, potential for the generation of important oscillations.

The comparison was conducted for several tests, including benchmark tests from the Convection–Diffusion Forum [16] and a real estuary (Tagus estuary, Portugal). Plume and flow conditions range from simple to very complex. The methods were compared based on a suite of error measures traditionally used for transport models [15, 16].

This paper is divided into five sections and one appendix, besides this introduction. Section 2 reviews and discusses the two classes of methods under comparison: FCTs and nonlinear filters. Section 3 describes the formulation of the base model, the details of the implementation of each method within the base model, and the two algorithms for mass conservation in FCTs. Section 4 describes the tests selected for the comparison. In Section 5, the various methods are compared on a property-by-property basis to foster the identification of their relative merits. Finally, Section 6 summarizes the findings of this research. The error measures are given in the Appendix.

2. BACKGROUND

2.1. *Flux-Corrected Methods*

Flux-corrected transport algorithms (FCTs) have been used extensively for over two decades to eliminate oscillations in finite difference schemes (e.g., [17–20]). Lohner *et al.* [5] adapted FCTs to finite elements, for multiple dimensional problems. Bermejo and Staniforth [21] implemented FCT concepts in a Eulerian–Lagrangian framework.

The basic concept of FCTs is to generate a positive, high-accuracy solution through a combination of high- and low-order methods. The high- and low-order methods may have distinct formulations [18] or the low-order method can be obtained by adding diffusion to the high-order solution [5]. The FCT solution, c^{corr}, can be generically obtained at each node as

$$c_i^{corr} = \gamma_i c_i^H + (1 - \gamma_i) c_i^L, \tag{1}$$

where c_i^H and c_i^L are the concentrations computed with the high- and the low-order methods at node i, respectively, and $\gamma_i \in [0, 1]$ is a weighting factor. Following a widely used approach in FCTs, the corrected concentration can be further restricted to the interval defined by the minimum and maximum concentrations in the elements that contain node i in the previous time step [5]. For ELM models, the limiting concentrations are the minimum and maximum concentrations at the element that contains the foot of the characteristic line of node i [9].

The use of a single value of γ in the whole domain preserves mass, provided that both c^H and c^L are mass conservative. However, this approach will either lead to excessive diffusion or fail to eliminate the oscillations. The optimization of γ at each node avoids these problems [22] but can potentially introduce mass errors [9]. *A posteriori* mass correction algorithms thus become necessary. Priestley [9] proposed an algorithm for global mass conservation by modifying the values of γ in a global sense. However, since the introduction of mass errors in the flux-correction procedure is local, a local mass correction algorithm appears to be more appropriate. A local mass correction algorithm is proposed in Section 3.

The main characteristic of FCT methods is the strict elimination of unphysical values. However, these methods require the computation of two numerical solutions at each time step (a "high-order" and a "low-order" solution), which may lead to substantial computational costs. Using a lumped mass matrix formulation as the positive, low-order method makes the computational cost associated with this method negligible.

Including source and sink terms or reactions, which are present for instance in geochemical models (e.g., [23]), may be difficult within the original FCT formulations. The lowest and highest concentrations allowed at each node in the new time step are limited by the minimum and maximum concentrations at the previous time step for the element that contains the foot of the characteristic line of that node. If there are mechanisms adding or removing mass from the system (sources or sinks) or transferring mass from one chemical constituent to another, the FCT limiting concepts are difficult, if not impossible, to implement. The modification proposed in Section 3 avoids this limitation by using user-specified bounds, related to the specific problem to be addressed.

2.2. *Nonlinear Filters*

Filters have been extensively used to remove noise and unphysical values, in particular in the analysis of field data time series [24]. Within the scope of the simulation of the transport

equation, the concept of nonlinear filtering for the removal of oscillations was proposed by Mahlman and Sinclair [25]. In spite of the promising results, this type of method has attracted little attention.

Within the context of structured grids, the nonlinear filter proposed by Engquist *et al.* [7] provided excellent results for 1D and 2D simulations [26, 27]. The basic concept of this filter is to eliminate local extrema by correcting the concentration at the corresponding node and at an adjacent node by the same amount.

Fortunato and Oliveira [28] extended this method to two-dimensional unstructured grids to filter bathymetric oscillations. This filter was further modified by Oliveira and Fortunato [8] to preserve local and global mass in depth-averaged transport simulations. The filter eliminates local extrema in the concentration field by equalizing the peak concentration and the concentration at the adjacent node that leads to the largest concentration difference. The concentration at both nodes is calculated by avoiding mass changes relative to the unfiltered solution.

Since excessive damping is a drawback often associated with the use of filters, Oliveira and Fortunato [8] proposed and explored local implementations of the 2D nonlinear filter. Given user-specified bounds on the concentration field, the filter is only applied at the nodes where these bounds are exceeded. An example is the propagation of salinity in an estuary, where the concentration cannot be negative or larger than the salinity at the ocean boundary, at least in the absence of evaporation. The concept of bounding the concentrations based on the problem at hand will be further explored in the following sections, for both nonlinear filters and FCT methods.

The nonlinear filter compared favorably with an implementation of the local use of a low-order method, providing both less damping and better mass conservation [8]. Therefore, the localized use of low-order methods was omitted in the present analysis. The filter also proved highly superior to increasing the diffusion coefficient [8].

Numerical filters have some advantages relative to other methods for the elimination of oscillations. Since they are applied *a posteriori*, they can be implemented in any model without any restriction on the formulation or on the problem at hand. Sources, sinks, and reactions are easily handled, and the filter does not need to be adapted for each specific problem. Filters are also computationally inexpensive, as they can be applied with large time steps, and only when oscillations exist. Finally, their judicious, localized application may limit the introduction of numerical diffusion.

3. DESCRIPTION OF THE NUMERICAL SCHEMES

This section presents the 11 numerical schemes used in the paper, including the base model and the 10 formulations for oscillation reduction. These 10 formulations have different approaches regarding the basic concept behind them (FCT, nonlinear filter, or a combination of the two), definition of FCT limiting bounds, existence and type of mass correction scheme for FCTs, and *a posteriori* oscillation removal for filters. The acronyms and general properties of these schemes are summarized in Table I.

3.1. *Base Model—VELA*

The base model (BM) used in this study, VELA, is a Eulerian–Lagrangian model, which combines finite volumes and finite elements with a high-accuracy integration based on

TABLE I
Definition of Formulations

Acronym	Filter	Flux corrected	Mass correction
BM	No	No	No
BM-F	Yes	No	Yes
BM-F-C	Yes (with clipping)	No	Yes
OFC	No	Yes (original)	No
MFC	No	Yes (modified)	No
GM-OFC	No	Yes (original)	Yes (global)
LM-OFC	No	Yes (original)	Yes (local)
GM-MFC	No	Yes (modified)	Yes (global)
LM-MFC	No	Yes (modified)	Yes (local)
GM-MFC-F	Yes	Yes (modified)	Yes (global)
LM-MFC-F	Yes	Yes (modified)	Yes (local)

subdivision quadrature [14]. The use of finite volumes for integration of the transport equation avoids local and global mass errors so often associated with ELMs [30], while the use of finite elements for domain and variable definition allows for flexible domain discretization and high accuracy. The use of integration techniques for the evaluation of the integrals at the feet of the characteristic lines was shown to be superior to the traditional interpolation methods [31].

As in all ELMs, the transport equation in VELA is split into two simpler equations, solved sequentially: advection is solved by the backward method of characteristics using an adaptive, embedded fourth-order Runge–Kutta method [29], while diffusion is solved using node-centered control-volume concepts. Linear finite elements are used both in geometry representation and in variable definition. Integration in time is performed with an α-method; i.e.,

$$\int_{\Omega_{cv}^{n+1}} c^{n+1}\, d\Omega_{cv}^{n+1} - \frac{\alpha \Delta t}{\bar{H}^{n+1}} \int_S \left(DH \frac{\partial c}{\partial x_k} \right) \vec{k} \Big|^{n+1} dS^{n+1}$$

$$- \int_{\Omega_{cv}^{n}} c^{\xi}\, d\Omega_{cv}^{n} - \frac{(1-\alpha)\Delta t}{\bar{H}^{\xi}} \int_S \left(DH \frac{\partial c}{\partial x_k} \right) \vec{k} \Big|^{\xi} dS^{n} = 0, \tag{2}$$

where ξ denotes the feet of the characteristic lines at instant n, D is the diffusion coefficient, Δt is the time step between time instants n and $n+1$, Ω_{cv}^{n} and Ω_{cv}^{n+1} are the surfaces of the control volume at instants n and $n+1$, $\bar{H}$ is the average depth in the control volume at each time instant, S_n and S_{n+1} are the boundaries of Ω_{cv}^{n} and Ω_{cv}^{n+1}, respectively, $\vec{k}$ is the outward unit normal on S, and α is the time discretization weight.

The evaluation of the integrals is performed after defining the control volume limited by the characteristic lines of the corners of the Voronoi polygon of each node (Fig. 1a). The area integrals at time $n+1$ are evaluated analytically in each quadrangle. The area integrals at the feet of the characteristic lines are evaluated by subdivision quadrature [14, 30]: each triangle formed by the foot of the nodal characteristic line and two sequential corners of the tracked image of the Voronoi polygon is divided into a user-specified number of subtriangles (Fig. 1b). Concentrations and depths are interpolated at each subtriangle corner and then assumed linear within each subtriangle.

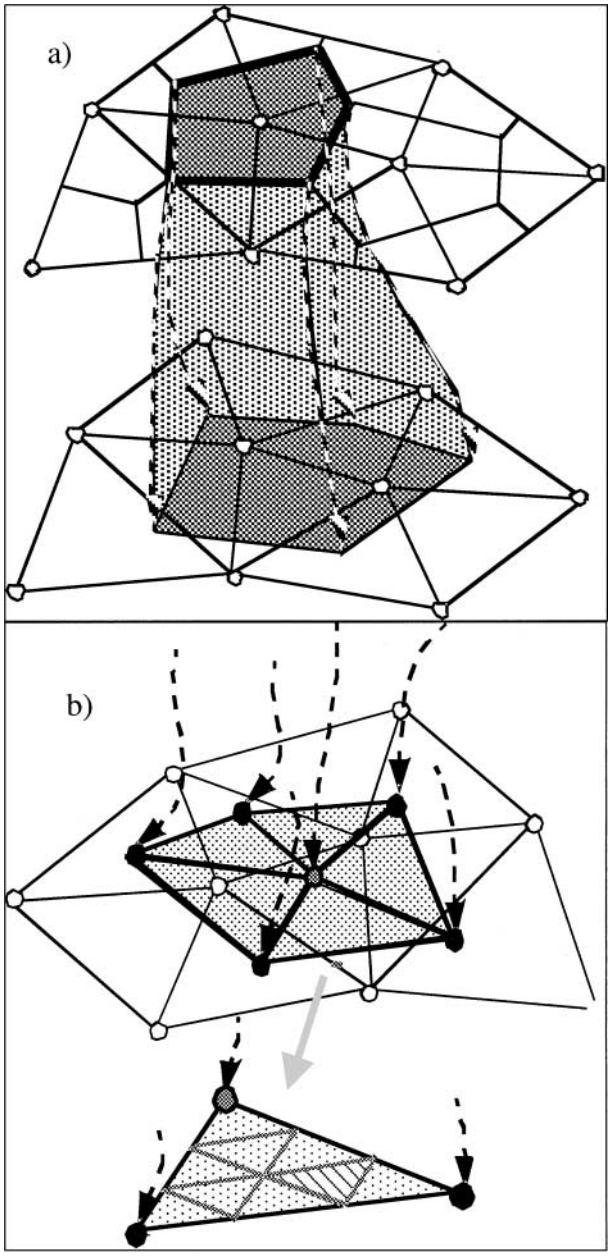

FIG. 1. (a) Definition of the control volume along the characteristic lines, between time steps $n + 1$ and n. (b) Region of integration at the feet of the characteristic lines, defined by the feet of the characteristic lines of the node and the corners of the Voronoi polygon. Definition of the integration triangles and their subdivision into a user-specified number of subtriangles.

3.2. *Nonlinear Filter*

The nonlinear filter formulation (BM-F) first computes an initial solution at the new time step using the BM. Local extrema in the concentration field are then eliminated, while mass is preserved locally. The algorithm can be summarized as follows.

1. Identification of the local extrema. A minimum occurs for $c_i < c_{min}$, where $c_{min} = \min\{c_j : j = 1, NEI\}$, and a maximum occurs for $c_i > c_{max}$, where $c_{max} = \max\{c_j : j = 1, NEI\}$. NEI is the number of nodes sharing an element with node i.

2. Identification of the neighboring node m to be used in the filtering procedure, for each extremum. For a local minimum, the node with c_{max} is used, while for a local maximum the node with c_{min} is used.

3. Adjustment of the concentrations at nodes i and m. These concentrations are set to be equal and mass changes in the set of elements that contain any of the two nodes are prevented; i.e.,

$$\Delta M = \int_A H(c - c^f)\, dA = 0 \tag{3}$$

where H is the total depth and the superscript f indicates a filtered value. The integration is performed over the elements containing any of the two nodes whose concentration is being adjusted (Fig. 2). Note that except for nodes i and m, the concentration is unchanged, which simplifies considerably the solution of (3).

For linear triangular elements, Eq. (3) becomes

$$\begin{aligned} c_i^f &= c_m^f = \left(-\frac{B_2}{B_1} c_m - c_i\right) \Big/ \left(-\frac{B_2}{B_1} - 1\right), \\ B_1 &= 2H_i \sum_{k=1}^{n\_el\_i} A_k + \sum_{k=1}^{n\_nd\_i} H_k \sum_{j}^{n\_el\_k\_i} A_j, \\ B_2 &= 2H_m \sum_{k=1}^{n\_el\_m} A_k + \sum_{k=1}^{n\_nd\_m} H_k \sum_{j}^{n\_el\_k\_m} A_j, \end{aligned} \tag{4}$$

where c_m and c_m^f are the concentrations at node m before and after the filter, c_i and c_i^f are the concentrations at node i before and after the filter, H is the nodal depth, A is the elemental area, n_el_i and n_el_m are the number of elements around nodes i and m, n_nd_i and n_nd_m are the number of nodes around nodes i and m, and $n\_el\_k\_i$ and $n\_el\_k\_m$ are number of elements that share nodes i and k and nodes m and k.

4. Steps 1 to 3 are repeated iteratively, at each time step, until all local extrema have been eliminated.

The application of the filter can be limited to nodal concentrations that exceed user-specified bounds. Often, the concentration is physically bounded by the conditions of the problem. For instance, salinity in an estuary should not be negative or exceed the ocean value used as a boundary condition (if the model is not considering evaporation). In this case, it is reasonable to limit the application of the filter to areas where a user-specified bound is being exceeded.

Although the application of the filter drastically reduces numerical oscillations, it may not fully eliminate them. Therefore, a clipping procedure was implemented, in which any nodal concentration that exceeds the user-specified bounds is set equal to these bounds. This procedure is denoted BM-F-C.

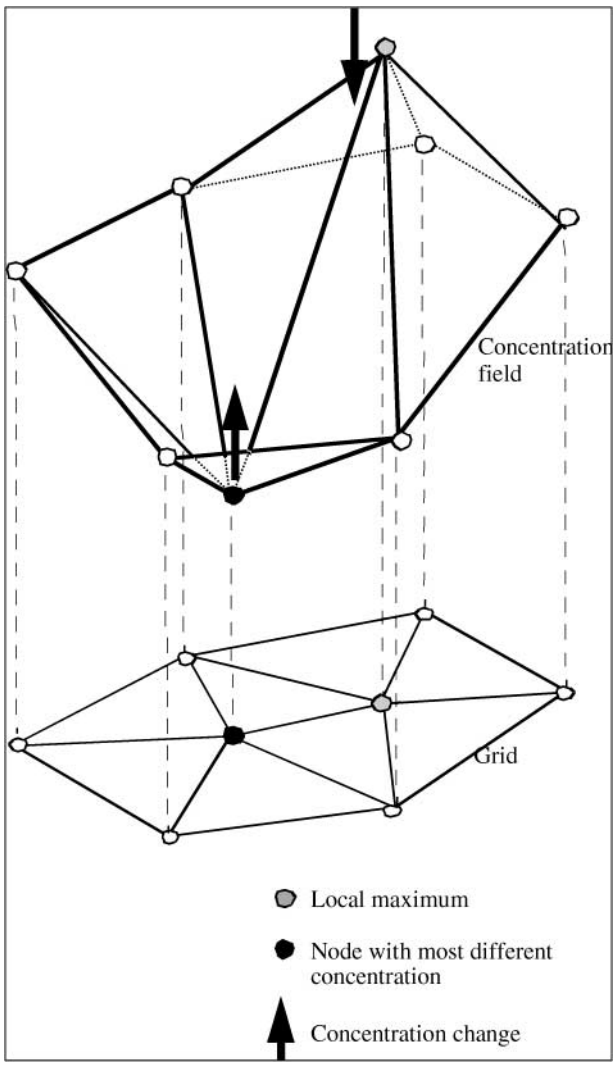

FIG. 2. Nonlinear filter: example of nodal selection and nodal movement.

3.3. *Flux-Corrected Schemes*

All flux-corrected transport schemes used in this paper are based on the generic definition (Eq. (1)). The low-order method is based on the BM, using a lumped mass matrix. This method is unconditionally positive while being less diffusive than other lumping schemes, as a result of the high-accuracy integration of the forcing vector. For the two methods described in this subsection, the original flux-corrected method (OFC) and the modified flux-corrected method (MFC), the BM is used in Eq. (1) as the high-order method.

In OFC, as in traditional FCT methods, the weighting factor, γ_i, is defined as the largest possible value at each node, given the restrictions

1. $\min(c_i^H, c_i^L) < c_i^{corr} < \max(c_i^H, c_i^L)$, i.e. $\gamma_i \in [0, 1]$,
2. $\min(c_i^\xi) < c_i^{corr} < \max(c_i^\xi)$,

where $\min(c_i^\xi)$, $\max(c_i^\xi)$ are the minimum and maximum concentrations in the element

that contains the foot of the characteristic line of node i at time n. The second restriction prevents the generation of new extrema in the concentration field, by limiting the corrected concentration to the range of values in the element that contains the foot of the characteristic line. However, this condition also limits the ability of the model to adapt to different grid refinements through mass redistribution. For complex models, involving sources, sinks, and reactions, the application of the second condition may be difficult, if not impossible.

These limitations motivated the development of a modified flux-corrected scheme (MFC). Rather than limiting the concentrations based on values in the previous time step, the MFC uses the concept of user-specified, problem-dependent, bounding limits: c^{upper}, c^{lower}. These limits can vary both in space and in time, depending on the specific problem. In this paper, the lower and upper bounds, when applicable, were specified as time- and space-independent values. The definition of the weighting factor for the MFC must thus obey the following restrictions:

1. $\min(c_i^H, c_i^L) < c_i^{corr} < \max(c_i^H, c_i^L)$, i.e. $\gamma_i \in [0, 1]$.
2. $c^{lower} < c_i^{corr} < c^{upper}$.

3.4. *Mass Correction Algorithms for Flux-Corrected Schemes*

The application of the FCT concepts, either as OFC or MFC, does not take into account the impact on mass conservation of the combination of the low- and high-order schemes through the weighting factor. In the framework of Eulerian–Lagrangian methods, Priestley [9] proposed a mass correction scheme that targets global mass conservation. The general concept of this scheme is to iteratively modify the weight of the high- and low-order solutions at each node by searching for the largest possible weights for the high-order method that maintain positivity and minimize global mass conservation errors.

This algorithm was adapted here for application to the control-volume finite element ELM concepts used in the BM. The algorithm used herein has the following steps.

1. Compute the integrated difference between high- and low-order solutions for each node, β_i, using the set of elements that contain node i; i.e.,

$$\beta_i = \int (c^H - c^L) H \, d\Omega_i, \tag{5}$$

where H is the total depth.

2. Compute c^*; i.e.,

$$c^* = \sum_{i=1}^{n\_nodes} \gamma_i \beta_i, \tag{6}$$

where n_nodes is the total number of nodes in the domain and γ_i are the weighting factors used in Eq. (1).

3. Targeting the mass conservation of the high-order solution, check whether the sign of c^* is correct:

$$\begin{gathered} \text{if } c^* > \int c^H H d\Omega - \int c^L H d\Omega, \quad \text{then } c^* = \int c^H H d\Omega - \int c^L H d\Omega; \\ \text{otherwise } c^* = \int c^L H d\Omega - \int c^H H d\Omega. \end{gathered} \tag{7}$$

4. Compute the indicator *iflag* and the initial estimates for the weighting factor γ^*:

$$\begin{aligned} &\text{if } \beta_i < 0 \begin{cases} iflag_i = 1, \\ \gamma_i^* = \gamma_i; \end{cases} \\ &\text{if } \beta_i >= 0 \begin{cases} iflag_i = 0, \\ \gamma_i^* = 0. \end{cases} \end{aligned} \tag{8}$$

5. Define average weighting factor, γ_{av}:

$$\gamma_{av} = \frac{c^* - \sum_{i=1}^{n\_nodes} \gamma_i^* \beta_i iflag_i}{\sum_{i=1}^{n\_nodes} \beta_i (1 - iflag_i)}. \tag{9}$$

6. If the best mass conservation was achieved, update the values of the weighting factor and recalculate the concentration:

$$\text{if } \gamma_{av} \leq \gamma_i, \quad \text{for all } iflag_i = 0 \Rightarrow \gamma_i = \gamma_{av}, \quad \text{for all } iflag_i = 0. \tag{10}$$

Then recalculate the concentration based on the new weighting factors and stop the mass correction procedure.

7. If the best mass conservation was not achieved, update values of weighting factor:

$$\begin{aligned} &\text{if } iflag_i = 0 \quad \text{and} \quad \alpha_{av} > \gamma_i \Rightarrow \gamma_i^* = \gamma_i; \\ &\text{if } iflag_i = 0 \quad \text{and} \quad \alpha_{av} \leq \gamma_i \Rightarrow \gamma_i^* = \gamma_{av}. \end{aligned} \tag{11}$$

8. Return to step 5 until convergence is reached.

The application of this procedure to the OFC and to the MFC generates the global mass correction original flux-corrected scheme (GM-OFC) and the global mass correction modified flux-corrected scheme (GM-MFC).

Since oscillations are generated in specific regions of the domain, we propose a new localized mass correction procedure. This algorithm first identifies the nodes where the high-order solution was replaced by a corrected concentration and uses one of the adjacent nodes to compensate for the mass imbalance introduced by the FCT. This algorithm is based on the same concepts applied for the mass conservation of the nonlinear filter.

For each node i where the high-order solution was not used, do the following.

1. Within the adjacent nodes (nodes that share one element with node i), find the minimum and maximum concentrations (c_{min}, c_{max}) and the difference between these concentrations and c_i^{corr}:

$$\begin{aligned} d_{min} &= abs\left(c_i^{corr} - c_{min}\right), \\ d_{max} &= abs\left(c_i^{corr} - c_{max}\right). \end{aligned} \tag{12}$$

2. Note that the node selected for the mass correction is the one that leads to the largest concentration difference (node m).

3. Correct the concentration at node m to maintain the mass of the high-order solution within the set of elements shared by node i and node m; i.e.,

$$\Delta M = \int_A H(c^H - c^{corr})\, dA = 0, \tag{13}$$

where H is the total depth and the superscript *corr* indicates a corrected value.

The integration is performed over the elements containing any of the two nodes whose concentration is being adjusted. Equation (14) leads to

$$\begin{aligned} c_m^{m\_corr} &= c_m + \frac{B_1}{B_2}\left(c_i^H - c_i^{corr}\right), \\ B_1 &= 2H_i \sum_{k=1}^{n\_el\_i} A_k + \sum_{k=1}^{n\_nd\_i} H_k \sum_{j}^{n\_el\_k\_i} A_j, \\ B_2 &= 2H_m \sum_{k=1}^{n\_el\_m} A_k + \sum_{k=1}^{n\_nd\_m} H_k \sum_{j}^{n\_el\_k\_m} A_j, \end{aligned} \tag{14}$$

where c_m and $c_m^{m\_corr}$ are the concentrations at node m before and after the mass correction, respectively, c_i^H and c_i^{corr} are the concentrations at node i for the high-order solution and after applying the FCT, H is the nodal depth, A is the elemental area, n_el_i and n_el_m are the number of elements around nodes i and m, and n_nd_i and n_nd_m are the number of nodes around nodes i and m, $n\_el\_k\_i$ and $n\_el\_k\_m$ are the number of elements that share nodes i and k and nodes m and k, respectively.

Application of the above algorithm to OFC and MFC generates the local mass correction FCTs (LM-OFC and LM-MFC). The concentration at the adjacent node m is also limited by the upper and lower bounds on the LM-MFC and by the concentrations at the element that contains the foot of the characteristic line of node m on the LM-OFC.

3.5. *Combination of Nonlinear Filter and Mass-Corrected Flux-Transport Schemes*

Two additional formulations are proposed based on a combination of the nonlinear filter and modified FCT concepts (GM-MFC-F and LM-MFC-F). The rationale behind these new formulations is the choice of a high-accuracy method where oscillations were already almost eliminated as the high-order method in a FCT general framework. These formulations are implemented by using the nonlinear filtered solution (BM-F) as the high-order method in the modified flux-corrected method, applying either a global mass correction (GM-MFC-F) or a local mass correction (LM-MFC-F).

4. EXPERIMENTAL SETUP

This section presents the four tests chosen for the comparison, which include both benchmark tests (rotating Gauss, rotating cylinder, and convection of concentration hills within closed streamlines) and the transport of a Gauss plume in the Tagus estuary forced by the results of a shallow-water model (ADCIRC [32]).

Plume conditions (Fig. 3) include smooth concentration gradients in the Gauss plume tests (rotating Gauss test and Tagus estuary) and sharp concentration gradients in the rotating

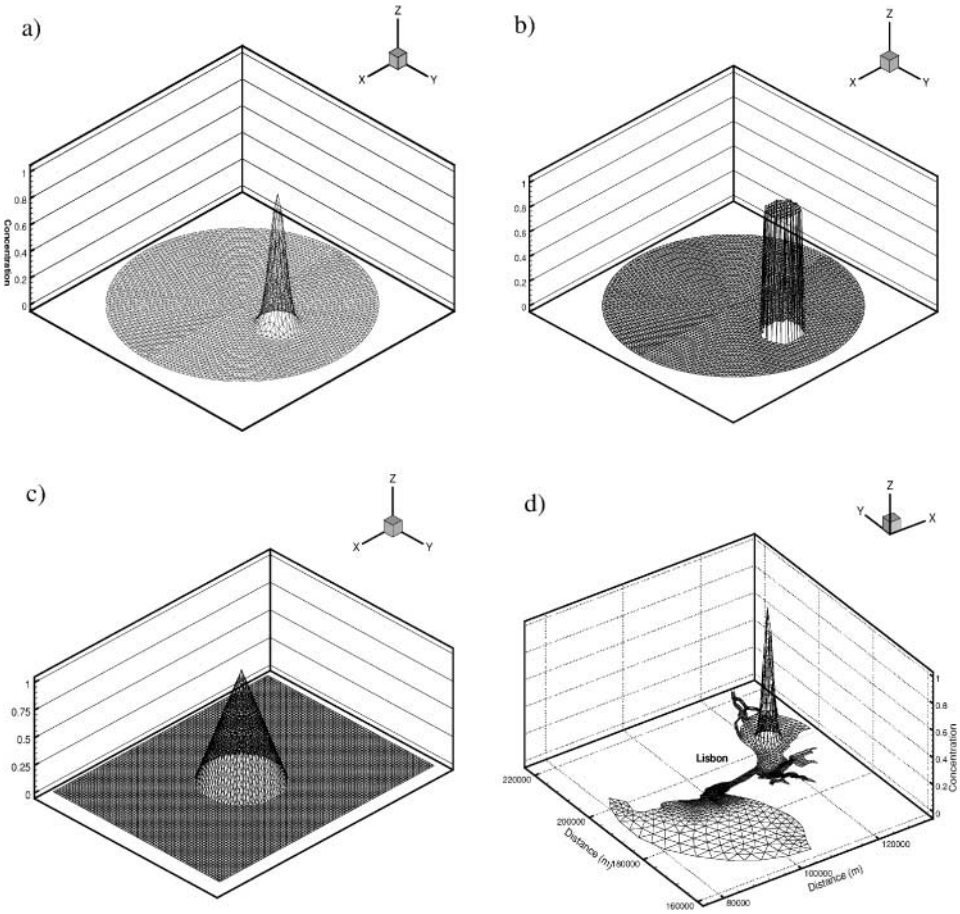

FIG. 3. Initial conditions: (a) rotating Gauss; (b) rotating cylinder; (c) CDF5; (d) Gauss in the Tagus estuary.

cylinder test and in the convection of concentration hills within closed streamlines test (Convection–Diffusion Forum test 5 (CDF5)).

Initial conditions for the rotating Gauss test and Gauss plume in the Tagus estuary are defined as

$$C_0 = \exp\left(\frac{(x - x_0)^2}{2\sigma_0^2} + \frac{(y - y_0)^2}{2\sigma_0^2}\right), \tag{15}$$

where x and y are the space coordinates, (x_0, y_0) is the initial location of the plume center, and σ_0 is the standard deviation of the initial Gauss hill.

For the CDF5 test, the initial concentrations are given by

$$C_0 = \begin{cases} 1 - \sqrt{\frac{(x-x_0)^2}{r_0^2} + \frac{(x-x_0)^2}{r_0^2}}, & \text{if } (x - x_0)^2 + (y - y_0)^2 \le r_0^2, \\ 0, & \text{otherwise}, \end{cases} \tag{16}$$

where r_0 is the cone base radius.

Finally, the initial conditions for the rotating cylinder are defined as

$$C_0 = \begin{cases} 1, & \text{if } (x - x_0)^2 + (y - y_0)^2 \le r_0^2, \\ 0, & \textit{otherwise}, \end{cases} \tag{17}$$

where r_0 is the cylinder base radius.

Forcing flow fields, which have an important impact on oscillation generation, range from simple in the rotating Gauss and rotating cylinder tests, to very complex in CDF5 and the Tagus estuary tests (Fig. 4). Indeed, the Gauss plume of the Tagus estuary test is forced by a fully nonlinear flow simulation. This simulation was conducted using the depth-averaged shallow-water model ADCIRC [32], forced by the main tidal constituent (M_2) and its major harmonics (M_4, M_6, and Z_0) at the ocean boundary, and by a constant river flow at the upstream boundary. This flow field is strongly space and time dependent (Fig. 4), due to the oscillatory nature of the ocean boundary forcing and the bathymetry and geometry of the Tagus. Further details of the tidal simulations for the full set of tidal constituents are given in Fortunato *et al.* [33, 34]. To avoid the propagation of flow mass errors to the transport simulations [13], the flow grid was obtained by refining the transport grid (Fig. 3) through division of all elements by 16.

For the rotating Gauss and rotating cylinder tests, the flow field is stationary and defined as

$$\begin{aligned} u &= -\omega y, \\ v &= \omega x, \end{aligned} \tag{18}$$

where ω is the angular velocity, taken as 2.09×10^{-3} s^{-1}.

For the CDF5, the flow field is also stationary:

$$\begin{aligned} u &= 0.32\pi \sin(4\pi x) \sin(4\pi y), \\ v &= 0.32\pi \cos(4\pi x) \cos(4\pi y). \end{aligned} \tag{19}$$

For the rotating plume tests, the time step and the number of time steps were selected to simulate a 360° rotation of the plumes. Additional details of all simulations are presented in Table II.

TABLE II
Test Parameters

Test	Rotating Gauss	Rotating cylinder	CDF5	Tagus estuary
Time step (s)	50	150	6	3600
Number of time steps	60	20	15	72
Standard deviation of Gauss plume/base radius (m)	200	200	1500	1000
Lower concentration bound	0	0	0	0
Upper concentration bound	—	1	—	—

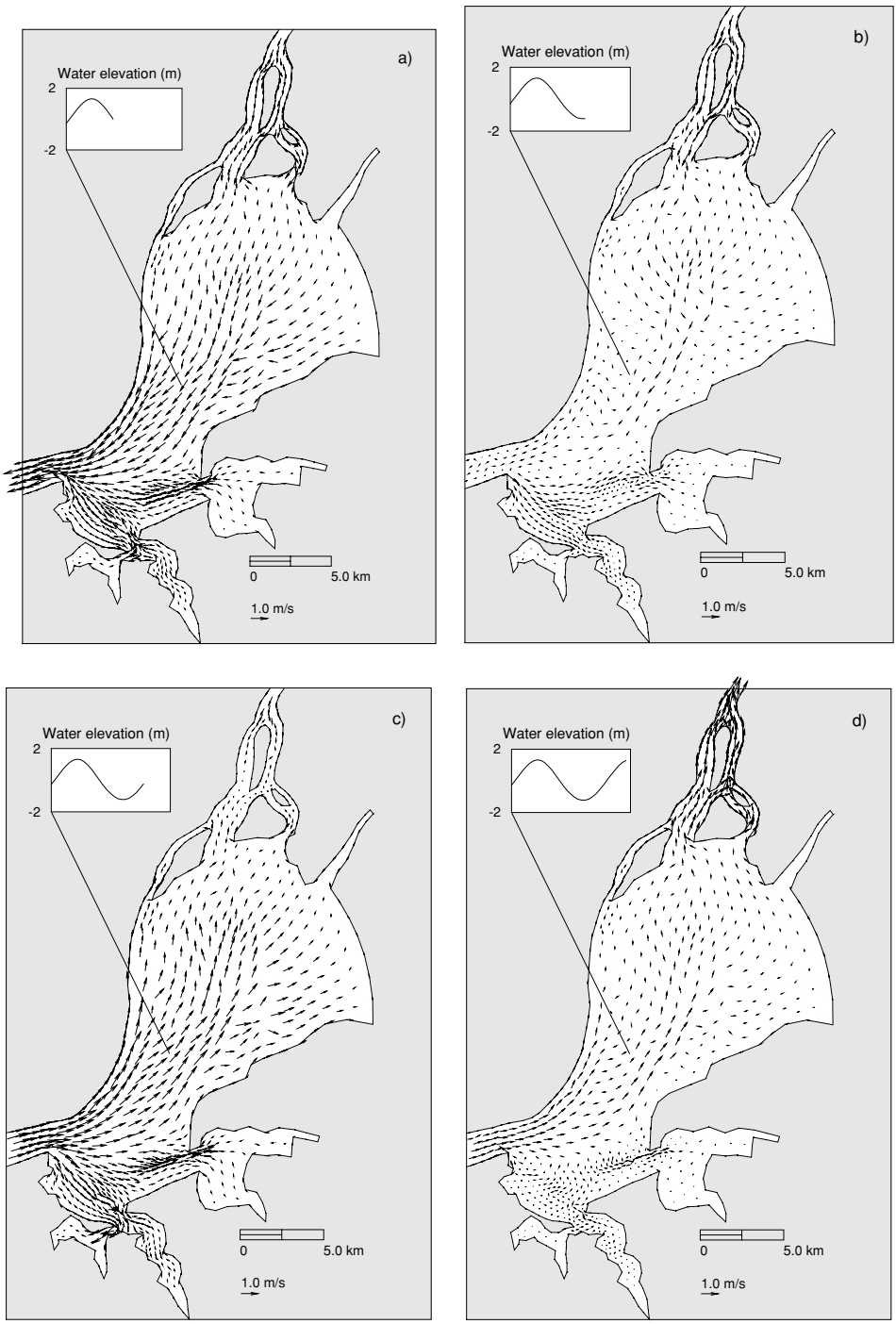

FIG. 4. Snapshots of the forcing flow field for the Tagus estuary test (velocities were interpolated into the transport grid for clarity): (a) ebb; (b) low-water tide; (c) flood; (d) high-water tide.

TABLE III
Measures of Numerical Oscillations

	Maximum negative concentration (−)		Maximum volume of negative concentration (m^3)	
Test	BM	BM-F	BM	BM-F
Rotating Gauss	-1.0×10^{-2}	-4.1×10^{-10}	-5.7×10^{3}	-6.4×10^{-5}
Rotating cylinder	-2.1×10^{-1}	-5.1×10^{-7}	-1.6×10^{5}	-4.3×10^{-3}
CDF5	-2.1×10^{-1}	-1.8×10^{-5}	-1.3×10^{5}	-5×10^{-1}
Tagus test	-1.0×10^{-1}	-9.6×10^{-8}	-1.5×10^{6}	-6.0×10^{-2}

5. RESULTS AND DISCUSSION

In this section, the 11 formulations are compared, based on the results from the tests described above. Since none of the methods emerged as optimal for all tests and all numerical properties, this comparison is made on a property-by-property basis, in order to identify the relative merits of each formulation.

5.1. *Numerical Oscillations*

The base model (BM) confirms its excellent results on all tests for all properties but numerical oscillations (Table III). The BM leads to considerable negative concentrations in particular for the tests with strong gradients (cylinder test and CDF5 tests, Fig. 3). In the presence of complex flow fields (Tagus and CDF5 tests), the BM also presents important oscillations, which can reach 10% of the peak concentrations (Tagus test). The volume of negative concentration (a measure of the "negative," unphysical mass) is also very large (Table III). The use of this otherwise very accurate method in real problems, involving simple transport or complex water-quality models, is thus limited by the large oscillations. Schemes that reduce or eliminate the oscillations are thus necessary.

All FCT-based formulations fully eliminate the oscillations that lead to negative concentrations or to concentrations above the cylinder top (Fig. 5). The nonlinear filter (BM-F) reduces the maximum negative concentrations by several orders of magnitude, leading to negligible negative concentrations in all tests (Table III). The oscillations that lead to concentrations above the cylinder top are fully eliminated. The volume of negative concentrations is also drastically reduced, between five and ten orders of magnitude (Table III).

In the CDF5, the performance of the nonlinear filter is worse than in the other tests. Due to the growth of concentration gradients in time, which increases the potential for oscillation generation (Fig. 5c), the nonlinear filter is less effective at removing the oscillations, leading to residual oscillations larger than in the other tests. These residual oscillations ($O(10^{-5})$), however, are still negligible (Table III).

Simple clipping of the values outside the bounds after the filter application (BM-F-C) can fully eliminate the remaining oscillations. However, this clipping can have a significant impact on mass conservation, as discussed below.

5.2. *Mass Conservation*

The base model preserves mass well in all tests. Mass conservation is only limited by the number of subdivisions used in the evaluation of the integrals at the feet of the characteristic

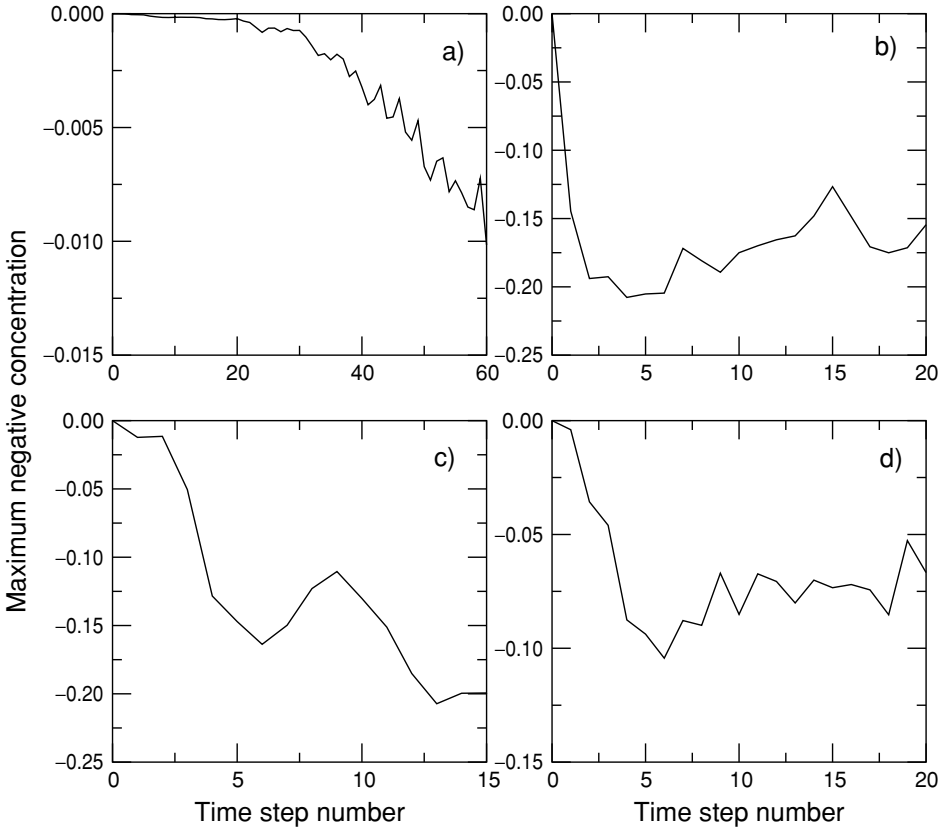

FIG. 5. Maximum negative concentrations for the base model: (a) rotating Gauss; (b) rotating cylinder; (c) CDF5; (d) Gauss in the Tagus estuary.

lines, and also by the flow mass balance in the Tagus estuary test [13]. Maximum relative mass errors vary between 8.0×10^{-5} and 1.6×10^{-2}.

All formulations for oscillation removal have an impact on mass conservation (Fig. 6). Both original and modified flux-corrected formulations (OFC and MFC) exhibit large mass errors because they change the concentration field without enforcing mass balance. The OFC has the largest errors in all tests, with relative errors up to 36% in the Tagus estuary. The modified formulation, MFC, reduces these errors consistently. Since limits on maximum nodal concentration are not imposed, the concentration field is able to adapt to the local refinement. Yet, mass errors remain large, since the major source of mass errors—the removal of the oscillations—was not compensated for in any way. The mass errors introduced by the flux-correction procedure (differences between errors for FCT methods and BM) are much larger for the tests with sharper concentration gradients, up to three orders of magnitude, than for the Gauss plume tests (rotating Gauss and Tagus estuary). This behavior was expected, as oscillations are generated in the presence of strong gradients.

The application of local or global conservation schemes on the OFC and MFC formulations are fundamental for acceptable mass conservation. For the rotating Gauss test, all formulations with mass correction procedures give similar mass errors (Fig. 6). This test has smooth concentration gradients and a simple flow field, so mass conservation does not limit the choice of the method for the oscillation removal. For the other tests, the local mass

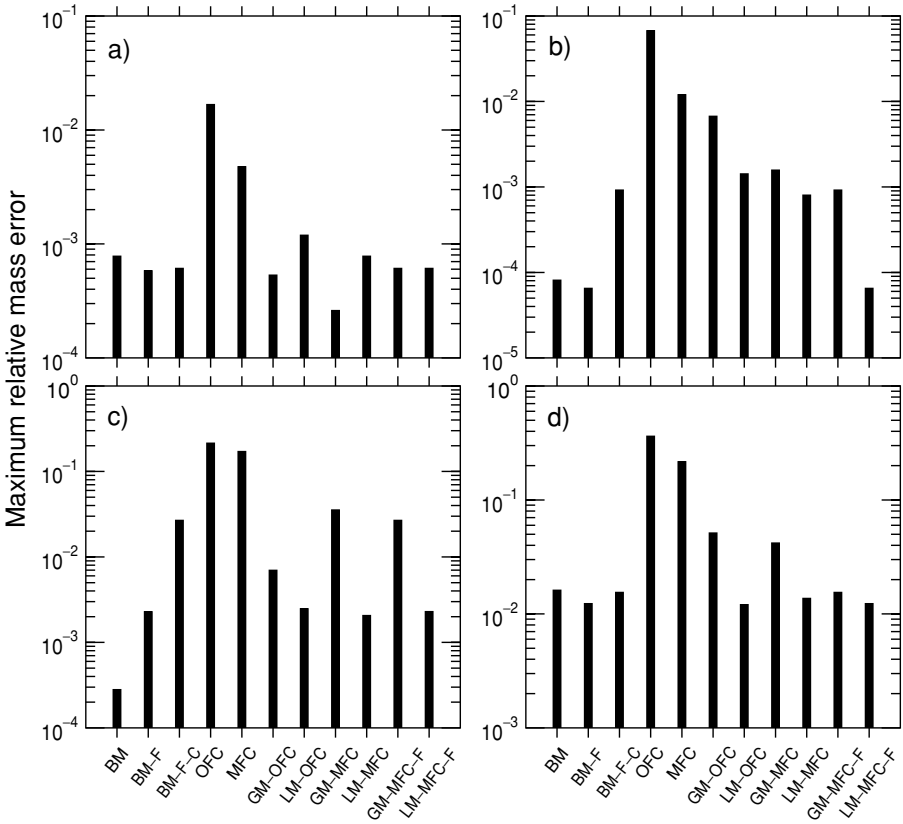

FIG. 6. Maximum relative mass errors: (a) rotating Gauss; (b) rotating cylinder; (c) CDF5; (d) Gauss in the Tagus estuary.

correction scheme leads to smaller mass errors than does the global mass correction scheme (up to one order of magnitude), both within the OFC and MFC.

The nonlinear filter (BM-F) generally leads to mass errors very similar to those of the base model. This behavior results from the mass correction procedure within this formulation that targets the mass of the base model locally.

The clipping of the oscillations after the application of the nonlinear filter (BM-F-C) has minor impact on mass conservation for the smoother concentration gradient tests (rotating Gauss and Gauss in the Tagus). However, for both the CDF5 and the rotating cylinder tests, clipping oscillations increases the maximum mass errors by an order of magnitude, due to the magnitude of the residual oscillations. This approach should thus be used with caution, guided by the specific problem to be solved.

Mass errors for the nonlinear filter without clipping are generally smaller than for the FCT methods using the BM as the high-order solution. These results motivated the development of the formulations based on a combination of the nonlinear filter and FCT concepts (GM-MFC-F and LM-MFC-F). The GM-MFC-F and LM-MFC-F lead to better mass conservation than the associated FCT methods using the BM as the high-order solution. The improved mass conservation results from the use of a high-order solution with much smaller oscillations, thus requiring a simpler mass correction after the application of the flux-corrected algorithm. The use of the local mass correction scheme emerges again as the best among FCT-based formulations.

From a mass conservation viewpoint, the nonlinear filter without clipping emerges as the best approach. FCT methods, with a local mass correction scheme, using either the BM or the nonlinear filter as the high-order solution, are also very attractive.

5.3. *Numerical Damping*

The diffusive properties of each formulation were evaluated through energy and peak losses. The conventional definition of energy errors [25] was modified to eliminate the impact of mass errors on energy loss (see the Appendix). This error measure was applied to all tests since it does not require an analytical solution. Peak errors require a maximum analytical concentration at each time step, so they were only evaluated for the first three tests.

The BM is by far the least diffusive method for all tests, with energy ratios very close to unity in most tests (Fig. 7) and very small peak errors for all tests but the rotating cylinder test (Fig. 8). The peak errors for the rotating cylinder arise from the oscillatory behavior of the BM in the presence of sharp concentration gradients. All alternatives for oscillation removal introduce numerical diffusion and peak loss.

Among the other formulations, the LM-MFC is the best for all tests on energy conservation (Fig. 7). This method is also the best on peak maintenance, except for the CDF5 test, where the MFC without any mass correction is slightly better (Fig. 8). The localized nature of

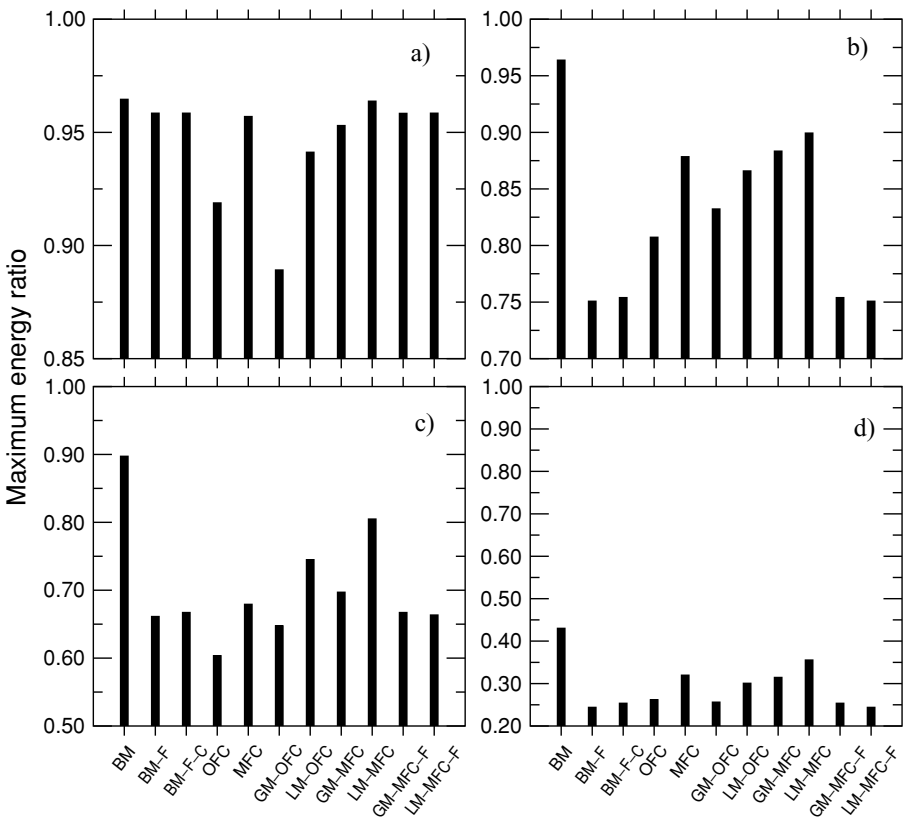

FIG. 7. Maximum energy ratios: (a) rotating Gauss; (b) rotating cylinder; (c) CDF5; (d) Gauss in the Tagus estuary.

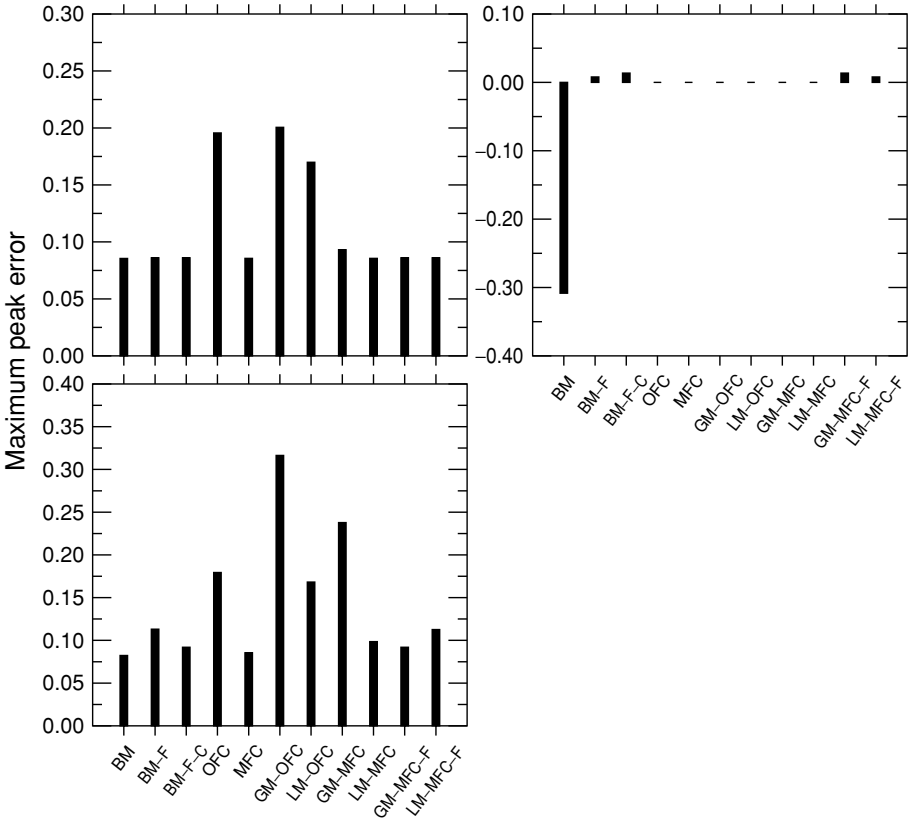

FIG. 8. Maximum peak errors: (a) rotating Gauss; (b) rotating cylinder; (c) CDF5.

the LM approach prevents the introduction of significant peak errors. The global mass correction, however, introduces significant peak errors both applied to MFC and applied to OFC.

Energy conservation is slightly better for the MFC than for the OFC method, both with and without mass corrections (Fig. 7). In general, mass correction algorithms will improve numerical damping by recovering lost mass. Global algorithms are less effective, as they may redistribute mass to the incorrect location.

All filter formulations introduce small peak errors (Fig. 8). These methods introduce, however, more numerical damping than FCTs for the sharp gradient tests. For the rotating Gauss test, filter formulations rate second best, with energy ratios very close to that of the LM-MFCs, as a result of the smooth concentration gradients. The relatively large numerical damping of the filter in the Tagus test is justified by the coarse discretization of the plume ($\sigma/\Delta x \approx 1.6$). Indeed, the mass correction algorithm within the filter shifts mass from the plume area to the adjacent nodes.

The analysis of the energy conservation and peak preservation showed that the algorithms for oscillation removal introduce some numerical diffusion. The modified flux-corrected formulation with local mass correction is the best choice, with minor peak loss and very small numerical damping on all tests. The nonlinear filter preserves peaks well but introduces significant numerical diffusion in the presence of sharp concentration gradients and for coarsely discretized plumes.

6. CONCLUSIONS

This paper presented a comparison of several flux-corrected transport methods and nonlinear filters for the elimination of numerical oscillations in finite element models. Within FCT methods, new formulations were proposed targeting mass conservation and added accuracy. All methods were implemented in a control-volume finite element ELM model [14] and tested in a suite of cases covering both benchmark tests and an estuarine application.

The elimination of numerical oscillations in highly accurate methods will in general generate mass errors and numerical diffusion. Among the methods analyzed, none emerged as the best for all tests and all error measures. The new modified flux-corrected method with a local mass correction scheme is a good approach for oscillation removal and leads to a generally small loss of accuracy relative to the original base model. Nonlinear filters, implemented both in a FCT concept and as a stand-alone formulation, provided comparatively good results. These formulations presented the best mass conservation properties and introduced little numerical diffusion, except in the presence of sharp gradients. Oscillations are not fully eliminated but are reduced by several orders of magnitude, down to negligible values.

Mass conservation was shown to be a major concern over the elimination of oscillations. While this property is accounted for within the nonlinear filters, traditional FCT formulations do not impose mass conservation. Explicitly recognizing the localized nature of numerical oscillations and, consequently, of the mass errors due to their elimination, we proposed a new method for local mass correction in FCTs. This method was shown to be superior to a previously proposed global mass correction scheme.

Within the selected base model, CPU times do not provide a basis for distinction among formulations. For the rotating Gauss test, CPU time increased less than 2% for all methods relative to the base model. This increase in CPU time is significantly smaller than the ones presented by Yost and Rao [27] for similar methods. The difference may be due to the use of larger time steps in ELMs, which implies fewer calls to the filter or the FCT algorythms.

Since nonlinear filters offer some advantages over FCT methods, such as unrestricted application to any problem or model, the choice of the algorithm should thus depend on the specific problem to be solved. For problems where mass conservation is vital (e.g., within geochemical models, where the transfer between chemical components is strongly controlled by mass), the nonlinear filter appears to be the best choice. In the presence of sharp gradients, the local mass correction FCT is the best option, in particular when numerical diffusion is a major concern.

APPENDIX: ERROR MEASURES

The comparison between the various methods is based on a suite of error measures that addresses the mass conservation, the oscillatory behavior, and the numerical diffusion. The definitions of the error measures used in the paper are given below. The superscripts *num* and *ex* indicate numerical and exact solutions, respectively. The subscripts *min* and *max* indicate minima and maxima, respectively.

1. The relative mass error, as a measure of the mass balance, is defined as

$$M(t) = \frac{1}{M_{init}} \int_{\Omega} H c^{num}(x, y, t)\, d\Omega, \tag{20}$$

where M_{init} is the mass at time step 0.

2. Dimensionless maximum negative concentrations, as a measure of the magnitude of the oscillations, are defined as

$$Neg(t) = \frac{c_{\min}^{num}(t)}{c_{\max}^{num}(t)}. \tag{21}$$

3. The volume of negative concentrations, as a measure of the mass of the negative concentrations, is defined as

$$NegVol(t) = \int_{\Omega} H \min(c^{num}, 0)\, d\Omega. \tag{22}$$

4. Peak error, as a measure of the reduction of the maximum concentrations, is

$$peak(t) = \left(c_{\max}^{ex}(t) - c_{\max}^{num}(t)\right) \Big/ c_{\max}^{ex}(t). \tag{23}$$

5. The relative energy error is an integral measure of the numerical damping. In order to eliminate the impact of mass errors in the energy errors, the definition of the energy errors [25] was modified to

$$E_m(t) = \frac{\int \left(c^{num} \dfrac{\int c^{ex}\, d\Omega}{\int c^{num}\, d\Omega} \right)^2 d\Omega}{\int (c^{ex})^2\, d\Omega}. \tag{24}$$

ACKNOWLEDGMENTS

This work was sponsored by the Fundação para a Ciência e a Tecnologia, project Protecção e Valorização da Zona Costeira Portuguesa, and by the Laboratório Nacional de Engenharia Civil, project Modelação Matemática de Processos Estuarinos e Costeiros. We thank R. A. Luettich, Jr., and W. W. Westerink for the model ADCIRC, A. M. Baptista for the software XMVIS used to produce Fig. 4, and three anonymous reviewers for their comments.

REFERENCES

1. P. Gresho and R. Lee, Don't supress the wiggles. They are telling you something! *Comput. Fluids* **9**, 223 (1981).
2. L. Demkowicz and J. T. Oden, An adaptive characteristic Petrov-Galerkin finite element method for convection-dominated linear and nonlinear parabolic problems in two space variables, *Comput. Methods Appl. Mech. Eng.* **55**, 63 (1986).
3. P. Zegeling, M. Borsboom, and J. van Kester, Adaptive moving grid solutions of a shallow-water transport model with steep vertical gradients, in *Proc. Comput. Methods in Water Resources XII*, edited by V. N. Burganos, G. P. Karatzas, A. C. Payatakes, C. A. Brebbia, W. G. Gray, and G. F. Pinder, Vol. 1, p. 427 (Computational Mechanics, Southampton, 1998).
4. A. B. Fortunato and A. Oliveira, An adaptive grid technique for the vertical structure of hydrodynamic models, *Appl. Math. Model.* **32**(8), 639 (1999).
5. R. Lohner, K. Morgan, J. Peraire, and M. Vahdati, Finite element flux-corrected transport (FEM-FCT) for the Euler and the Navier-Stokes equations, *Int. J. Numer. Methods Fluids* **7**, 1093 (1987).
6. J. R. Cheng, H. P. Cheng, and G. T. Yeh, A Lagrangian-Eulerian method with adaptively local zooming and peak/valley capturing approach to solve two-dimensional advection-diffusion transport equations, *Int. J. Numer. Methods Eng.* **39**, 987 (1996).

7. B. Engquist, P. Lotstedt, and B. Sjogreen, Non-linear filters for efficient shock computation, *Math. Comput.* **52**, 509 (1989).
8. A. Oliveira and A. B. Fortunato, Non-linear filtering of numerical oscillations in two-dimensional unstructured grids, in *Proc. Computational Methods in Water Resources XIII*, edited by L. R. Bentley, J. F. Sykes, C. A. Brebbia, W. G. Gray, and G. F. Pinder, Vol. 1, p. 381 (Balkema, Rotterdam, 2000).
9. A. Priestley, A quasi-conservative version of the semi-Lagrangian advection scheme, *Mon. Weather Rev.* **121**, 621 (1993).
10. A. Baptista, *Solution of Advection-Dominated Transport by Eulerian-Lagrangian Methods Using the Backward Method of Characteristics*, Ph.D. dissertation (Mass. Inst. Technol, Cambridge, MA, 1987).
11. M. A. Celia, T. F. Russell, I. Herrera, and R. E. Ewing, An Eulerian-Lagrangian localized adjoint method for the advection-diffusion equation, *Adv. Water Resour.* **13**(4), 187 (1990).
12. A. Oliveira and A. M. Baptista, On the role of tracking on Eulerian-Lagrangian solutions of the transport equation, *Adv. Water Resour.* **21**(7), 539 (1998).
13. A. Oliveira, A. B. Fortunato, and A. M. Baptista, Mass balance in Eulerian-Lagrangian transport simulations in estuaries, *J. Hydraul. Eng.* **126**(8), 605 (2000).
14. A. Oliveira, *Eulerian-Lagrangian Analysis of Transport and Residence Times in Estuaries and Coasts*, Ph.D. dissertation (Oregon Graduate Institute of Science and Technology, Portland, OR, 1997).
15. W. P. Budgell and M. D. Skogen, *Advection Schemes for Coastal Ocean Models on Unstrutured Triangular Meshes* (Fisken og Havet, Inst. of Marine Research, Bergen, Norway, 2000).
16. A. M. Baptista, E. E. Adams, and P. Gresho, Benchmarks for the transport equation: the Convection-Diffusion Forum and beyond, *AGU Coast. Estuarine Stud.* **47**, 241 (1995).
17. J. P. Boris and D. L. Book, Flux-corrected transport. I. SHASTA, a fluid transport algorithm that works, *J. Comput. Phys.* **11**, 38 (1973).
18. S. T. Zalesak, Fully multidimensional flux-corrected transport algorithms for fluids, *J. Comput. Phys.* **31**, 335 (1979).
19. R. G. Hills, K. A. Fisher, M. R. Kirkland, and P. J. Wierenga, Application of flux-corrected transport to the Las Cruces Trench site, *Water Resour. Res.* **30**(8), 2377 (1994).
20. S. A. Yost and P. M. S. V. Rao, Flux-corrected transport technique for open channel flow, *Int. J. Numer. Methods Fluids* **29**, 951 (1999).
21. R. Bermejo and A. Staniforth, The conversion of semi-Lagrangian advection schemes to quasi-monotone schemes, *Mon. Weather Rev.* **120**, 2622 (1992).
22. G. E. Georghiou, R. Morrow, and A. C. Metaxas, An improved finite-element flux-corrected transport algorithm, *J. Comput. Phys.* **148**, 605 (1999).
23. T. Wood and A. M. Baptista, A model for diagnostic analysis of estuarine geochemistry, *Water Resour. Res.* **29**(1), 51 (1993).
24. M. G. G. Foreman, W. R. Crawford, and R. F. Marsden, De-tiding: Theory and practice, *AGU Coast. Estuarine Stud.* **47**, 203 (1995).
25. J. D. Mahlman and R. W. Sinclair, Tests of various numerical algorithms applied to a simple trace constituent air transport problem, in *Fate of Pollutants in the Air and Water Environments*, edited by I. Suffet, Advances in Environmental Science and Technology, Vol. 8, p. 223 (Wiley–Interscience, New York, 1973).
26. W. Shyy, H.-H. Chen, R. Mittal, and H. S. Udaykumar, On the suppression of numerical oscillations using a non-linear filter, *J. Comput. Phys.* **102**, 49 (1992).
27. S. A. Yost and P. Rao, A non-linear filter for one- and two-dimensional open channel flows with shocks, *Adv. Water Resour.* **24**, 187 (2001).
28. A. B. Fortunato and A. Oliveira, On the representation of bathymetry by unstrutured grids, in *Proc. Computer Methods in Water Resources XIII*, edited by L. R. Bentley, J. F. Sykes, C. A. Brebbia, W. G. Gray, and G. F. Pinder, Vol. 2, p. 889 (Balkema, Rotterdam, 2000).
29. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran* (Cambridge Univ. Press, Cambridge, UK, 1992).
30. R. W. Healy and T. F. Russell, A finite-volume Eulerian-Lagrangian localized adjoint method for the solution of the advection-dispersion equation, *Water Resour. Res.* **29**(7), 2399 (1993).

31. A. Oliveira and A. M. Baptista, A comparison of integration and interpolation Eulerian-Lagrangian solutions of the transport equation, *Int. J. Numer. Methods Fluids* **21**(3), 183 (1995).
32. R. A. Luettich, Jr., J. J. Westerink, and N.W. Sheffner, *ADCIRC: An Advanced Three-Dimensional Circulation Model for Shelves, Coasts and Estuaries. Report 1: Theory and Methodology of ADCIRC-2DDI and ADCIRC-3DL* (Dep. of the Army, US Army Corps of Engineers, Washington, DC 1991).
33. A. B. Fortunato, A. M. Baptista, and R. A. Luettich, Jr., A three-dimensional model of tidal currents in the mouth of the Tagus estuary (Portugal), *Continent. Shelf Res.* **17**(14), 1689 (1997).
34. A. B. Fortunato, A. Oliveira, and A. M. Baptista, On the effect of tidal flats on the hydrodynamics of the Tagus Estuary, *Oceanologica Acta* **22**(1), 1 (1999).