Reviewer 1

General Comments

This work is a detailed study of the reproduction of a river plume by some state-of-the-art unstructured mesh models. The problem is introduced with an excellent analysis of the analytical solution and the results show the reproduction of various characteristics of the plume by the models. The use of different numerical schemes and other model parameters is discussed. The paper presents a considerable amount of work of very good quality and of great interest. However, although the first part is well written, the second part with results and discussion needs a substantial revision. Below are the main comments, which refer from section 5 onwards.

Dear Reviewer 1,

Thank you for your effort to review the paper and your insightful comments!

- This part of the paper is written as a technical report. The authors speak to a reader interested in reproducing their experiments. This greatly limits the paper and makes it less useful for readers interested in applying a model in a real situation. In particular, the paper should answer questions such as: what is the best numerical scheme for reproducing a river plume? What are the minimum horizontal and vertical resolutions still good? I would suggest adding a section of conclusions after section 6, which answers these questions, those posed in the Introduction (p.3, r80-83) and discusses the last sentence of the abstract;

  Thank you for the comment! Yes, we agree that the focus was largely on the reproduction of the test case, which still remains the main goal. However, in the revised manuscript we tried to avoid too technical formulations of the results, proposing a broader view on them. In the revised manuscript, we summarised our findings in the Conclusion Section as recommended. The manuscript aims at assessing the level of numerical mixing for different numerical schemes used in our runs, and also on documenting the test case. However, we cannot state which scheme would perform best in realistic cases, because many other factors may contribute in addition to numerical mixing. This is a much broader topic. We are, of course, interested in it and plan to address it in the future. We did not concentrate on the question about minimum/optimal horizontal/vertical resolution; to answer it, a new series of sensitivity runs should be presented and discussed, which is once again a subject of future work. However, we did some additional experiments: we fixed the depth to 10 meters everywhere (in the presented setups there is a slope, maximum depth is 30 m). This step provides increased resolution and supposes the usage of the layers with nearly constant depths. In this case all solutions are much closer to the analytical one. However, the test case is designed to stress the difference in the performance rather than to find the conditions when the analytical solution can be reached.

We have now added a Conclusion Section, which concentrates on the questions raised in the Introduction and in the last sentence of the Abstract:

As expected, all models reproduce the four prototypical zones shown in Figure 1, preserve freshwater volume in the system, and are stable. However, substantially different dynamics in each zone can be obtained using the same model but with different transport schemes. The major result of our comparative study is that accuracy in reproducing the analytical solution depends less on the applied model discretization or grid type than on the advection scheme. Table 3 illustrates how the choice of an advection scheme and limiter affects the performance. All runs listed there were carried out with the same number of vertical sigma layers and the same turbulence closure for vertical viscosity (eddy diffusivity is set to zero). Table 3 presents two final characteristics: the maximum offshore bulge spreading and total salinity fluxes after two inertial periods, enabling a better evaluation of the simulation. The total salinity fluxes characterize the net level of numerical diffusion in the system, or how closely the system adheres to the expected two-layer system with only two salinity classes. The smaller the absolute value of salinity fluxes, the lower the level of the numerical mixing in the considered simulation.

In runs where eddy vertical viscosity is turned on, the level of numerical diffusion and simulated plume characteristics are interrelated (Table 3, Section 5). If numerical mixing is larger, we get a smaller bulge offshore extent, a thicker plume, a bulge center closer to the coast, and a larger coastal current discharge. However, although they are related, they complement each other: the bulge spreading characterizes largely the horizontal part of dynamics, whereas total salinity fluxes characterize the vertical. When eddy viscosity is replaced by a constant background value (Subsections 6.2, 6.3), horizontal and vertical dynamics become less dependent.

The order of accuracy of a given scheme is as important as the type of limiter. The fct1, fct2, geometrical and Superbee limiters outperformed fct2, fct3 and Sweby's limiters in a sense of two considered characteristics (Table 3). In general, the Superbee limiter, with its anti-diffusive properties can be marked as a best limiter and Sweby's as the worst (however, only in vertical dimension) for the considered task. Among the considered advection schemes, the best performer was a hybrid MUSCL-type advection scheme (3d-4th order). Combined with the fct1 limiter, it gives the best result for the runs for which eddy vertical viscosity was calculated based on turbulence closure. However, for a highly accurate advection solution, noise may appear, attributed to the discretization type (or of other origin). Among the tested second-order advection schemes, the Miura and TVD schemes performed better than the upwind scheme.

In realistic cases, many other factors, such as physical mixing, influence the plume dynamics in addition to the numerical mixing. We therefore refrain from selecting the best scheme based solely on the results presented above. However, the knowledge of the extent the numerical mixing may affect the solution is an important preliminary step, and we hope that the approach of this paper will be of interest to other modelers.

New Table 3. The description of runs and their results in respect to changes of the advection scheme and limiter types. The discharge in all setups was equal to 3000 m^3/s, 40 vertical layers were used, eddy viscosity coefficient was calculated based on second-order turbulence model (k-ε style). ('FV' - Finite Volume, 'DG-FEM' - Discontinuous Galerkin Finite Element Method.)

<table>
<tr><th>N of run</th><th>Adv. scheme</th><th>Limiter</th><th>Model/grid/ discretization</th><th>Bulge max. offshore spreading, 35h, km</th><th>Theoretical prediction (lab. studies), 20-35h, km</th><th>Total salinity fluxes, psu m/s</th></tr>
<tr><td>1</td><td>85% of 3rd order + 15% of 4th order</td><td>fct1</td><td>FESOM-C/tri/ FV quasi-B-grid</td><td>19.9</td><td rowspan="9">24</td><td>-5</td></tr>
<tr><td>2</td><td>2nd order (upwind)</td><td>geom.</td><td>Thetis/tri/ DG-FEM</td><td>16.2</td><td>-6.3</td></tr>
<tr><td>3</td><td>85% of 3rd order + 15% of 4th order</td><td>fct2</td><td>FESOM-C/tri/ FV quasi-B-grid</td><td>20.1</td><td>-5</td></tr>
<tr><td>4</td><td>85% of 3rd order + 15% of 4th order</td><td>fct3</td><td>FESOM-C/tri/ FV quasi-B-grid</td><td>16.1</td><td>-6.2</td></tr>
<tr><td>5</td><td>2nd order (Miura)</td><td>fct1</td><td>FESOM-C/tri/ FV quasi-B-grid</td><td>17.3</td><td>-5.8</td></tr>
<tr><td>6</td><td>2nd order (upwind)</td><td>no</td><td>FESOM-C/tri/ FV quasi-B-grid</td><td>16.5</td><td>-13.1</td></tr>
<tr><td>11</td><td>85% of 3rd order + 15% of 4th order</td><td>fct1</td><td>FESOM-C/quad/ FV quasi-B-grid</td><td>19.7</td><td>-4.4</td></tr>
<tr><td>12</td><td>2nd-order (TVD)</td><td>superbee</td><td>GETM/quad/ FV C-grid</td><td>17.0</td><td>-4.6</td></tr>
<tr><td>13</td><td>3d order HSIMT (TVD)</td><td>Sweby's</td><td>GETM/quad/ FV C-grid</td><td>17.0</td><td>-16.7</td></tr>
</table>

- The part of the results is too long and difficult to read, it should be reduced where possible. Furthermore, after section 4, the English must be carefully checked and improved (you could contact a native English speaker), trying to use shorter sentences, better use of punctuation and to extend the explanation of some parts with complex concepts which, sometimes, are sketched out;

Thank you for the comment. We went through the Results section, removed some parts and rewrote the others. Also, the proofreading of the second part of the manuscript has been done according to the suggestion.

- I would move section 7 to the appendix, trying to use some tables. I would finish the paper, in a more traditional way, with the Conclusions.

  We put Section 7 into Appendix. The table with final characteristics is now in the Conclusion Section.

Specific Comments

- Throughout the paper, references should be made to the numbers of the sections, not to their name;

  Done.

- Table 1 would be more convenient at the beginning of section 5, where it is cited many times;

  Thank you, we put in the beginning of Section 5.

- From section 5 the Authors use "second (first) inertial period", which is a bit misleading. I would use "two (one) inertial periods" or "two rotational periods", in accordance with the first part of the paper;

  We have replaced first/second by one/two. Only in places, where there are ‘within’ or ‘over’, we use ‘second’ or ‘first’.

- I think that the comparison with analytical results and laboratory studies should be used more, both in the text and in the figures. In the figures, it would be useful to see these quantities. In any case, I leave the decision to the authors;

  In the text we emphasize the comparison to analytical solution more. We also added Table 3 (see previous answers).

- The figures with the vertical profiles have the x-axis inverted. I find this unintuitive; anyway, it is not so important;

  Thank you, we have modified all figures accordingly.

- p21r465-467: Explain more;

  In the revised manuscript we have deleted this piece of text.

- Fig. 6: Explain the various panels more. A line Fr = 1 would be useful;

  Done.

- p23r478-479: Explain more;

  Thank you. Done:

  The ratio between the length (along-shelf spread) and width (offshore spread) of the bulge called ellipticity (Avicola and Huq, 2003) is another parameter, which indicates the presence of numerical mixing in the system. Generally, numerical mixing tends to reduce the bulge external radius due to a decreasing salinity gradient (horizontal, vertical or both) in the near-field or bulge zone and the resulting reduction in plume-associated offshore velocities. Numerical mixing leads to a deepening of the bulge or/and to a changed angle of impingement, such that the center of the bulge gets closer to the coast: the bulge ends up being sliced off by the coastal wall. Numerical mixing therefore tends to increase the ellipticity. It thus comes at no surprise that in all triangular-mesh configurations, including run1, the ratio is too large compared to the expected number (Table 2).

- p24r495-500: Explain better;

  Ok, done:

  The position of the front of the coastal current can also provide a qualitative estimate of the level of numerical diffusion. Numerical mixing moves the bulge center closer to the coast, and hence a larger portion of freshwater enters the coastal current. The position of the head of the coastal current, or the magnitude of its discharge (compared to the analytical solution), can be used to diagnose numerical diffusion in the system (Fig. 5). Note that numerical diffusion levels may be even higher if the same small, offshore restriction of the bulge parallels a weakly developed coastal current. In such a case, the bulge and the coastal current are excessively thick (see e.g., run 6, Fig. 5).

- Section 5.3: Some parts of the text are missing (p27r516). Another error in r522. The text describes the differences in a concise way, with short comments. Sometimes it is difficult to understand;

  Yes, thank you, here is a mistake. Done!

- Section 5.4: Like the previous one, it is hard to understand. Use shorter sentences and describe better the methodology;

  Done.

- p29r557: not clear;

  We have added an explanation:

  ...Among all runs, runs 11 and 12 yield maximum freshwater volumes for the first salinity class. The released freshwater thus largely stays fresh and does not replenish intermediate salinity classes in runs 11 and 12 to the same extent as in the other runs. Runs 1 and 11 suggest that quasi-B discretization on the triangular grid can be a noticeable source of numerical mixing unless noise is suppressed by a filter.

- p30r563-564: explain better; p30r569: why? Explain more;

  Done for both comments:

  *It is known (e.g., Hetland, 2005; Burchard, 2020) that a larger isohaline area means less mixing for a given level of freshwater discharge through the isohaline. This fact is illustrated in Equation 16: $f(S) = F(S)/A(S)$. If we keep the diahaline freshwater discharge (**total** transport) F(S) constant and modify the isohaline area, which is situated in the denominator, the respective average diahaline freshwater flux, f(S), would increase. Increased diffusive fluxes in turn mean increased mixing, which in our case is purely numerical. Also the total volume of a particular class can be the same while mixing levels differ from case to case due to different isohaline areas (Eq. 14-16). The volume of the first class can also be larger from one case to another in line with a higher numerical diffusivity level; this would apply, for example, where a plume spreads relatively little and remains relatively thick. To avoid a wrong interpretation of the salinity-volume diagram, one should consider the area of the isohalines (Fig. 10) for each different run. Figure 10 depicts the isohaline areas for different salinity classes. Except for run 7 in Figure 10, which is characterized by the presence of physical eddy diffusivity, all curves have very similar shapes. These shapes reflect that a two-layer system is being considered. The layer occupied by the plume is characterized by low salinities; it is not completely fresh due to the presence of numerical mixing. The plume layer can be characterized by the rapidly growing part of the curve.*

- Fig.11 The curves are different in the panels, use the legends in each panel;

  Done.

- p32 605-615 Not clear, write better;

  Yes, thank you. Done:

  Figure 11b shows the transport through different isohalines: freshwater discharge through different isohalines (Fig. 11a) divided by the corresponding isohaline areas (Fig. 10). Transport naturally tends to decrease as the considered isohaline increases; but even so, the discharge can be the same (see, e.g., run 7 or run 13 in Fig. 10, 11a). Figure 11b sheds some light on vertical near-surface dynamics. For example, Figure 11b shows that run 2 has a large cross-isohaline transport for low salinities. Therefore, despite its generally good performance, this run is characterized by a blurry surface layer (Fig. 3). We also see the principal difference between runs 7 and 13: transport through the freshwater isohaline is largest for 13 but then rapidly decreases, and there is no pronounced mixed layer as there is in run 7.

- p34r650: compared to? Run 1?

  Compared to run 5. We have re-written the sentences.

- p35r669: First explain the purpose and then describe the runs.

  We have re-written the sentence.

- p36r688: Remove the text in brackets, it is not clear;

  Done.

- p37r690: Where? Explain better where the reader should look;

  Done.

- p38r726: Rephrase the sentence describing your findings not your suggestion.

  Done.

Technical Corrections

- p3r74: the

  Done.

- fig1: r is r_0?

  r is correct.

- p5r131: brackish -> fresh

We have been replaced.

- p5r142: we recommend.. -> we increased ... Describe the set-up, don't give recommendations.

Done.

- p5r146: we suggest -> What did you use? As before, describe the set-up.

Done.

- p7r158: below. Specify the section number.

Done.

- p9r200: a

Done.

- p12r302-305: u_b and h_0 are defined in sec2.1, say it somewhere or remember their definitions;

We have restated their meanings in the text

- p14r357: angel…

Fixed:)

- p21r464: typo

Done.

- p23r490: there

Done.

- Fig. 7: surface v-component?

Yes!, done, thank you.

- p24r505: Fig3, which panel?

We have specified the panel - Fig 3d.

- p28r547: intervening? I don't understand this sentence;

We rephrased the sentence, thank you.

- p23r550:.) ->).

Done.

- p30r559 signalizing?

It is rephrased.

- p32r591 than

‘Than’ is removed.

- p39r752 general

Thank you, removed.

Reviewer 2

Authors have done a lot of experiments, and some of them should be useful. While some of them duplicated with previous plume modeling, such as Xia's Cape Fear River Modeling (North Carolina State University, 2007) to discuss how the numerical scheme to impact plume's structure modeling, which is the key part of this manuscript. The horizontal resolution and vertical layers were given discussion in Xia et al., 2007 and other papers which were listed below. Numerical mixing can be considered as numerical errors, and authors need be carefully to utilize the error for the physical explanation. Please re-format this work and do a more than major revision to provide insight the plume community. Please remove most experiments which was conducted from these papers below and other literatures, make this manuscript concise and useful. I will provide more comments after new version, not this long, wordy unclear one. Also did authors simulate the internal wave (line 68?)

Niu, Q., Xia, M. (2021) "The behavior and wind-driven dispersions of two dynamically distinctive limnetic river plumes in a semi-enclosed basin," Estuarine, Coastal and Shelf Sciences.(In press)

Niu, Q., Xia, M., Ludsin, S.A., Chu, P.Y., Mason, D.M., Rutherford, E.S. (2018). "Highâ□□urbidity events in Western Lake Erie during ice-free cycles: Contributions of river-loaded vs. resuspended sediments," Limnology and Oceanography,00, 1-18.

Jiang, L., & Xia, M. (2016). "Dynamics of the Chesapeake Bay outflow plume: Realistic plume simulations and its seasonal, interannual variability," Journal of Geophysical Research: Oceans, 121, 1424-1445.

Xia, M., Xie, L., Pietrafesa, L.J., Whitney, M.M. (2011). "The ideal response of a Gulf of Mexico estuary plume to wind forcing: Its connection with salt flux and a Lagrangian view," Journal of Geophysical Research, 116, C08035.

Xia, M., Xie, L., Pietrafesa, L.J. (2010). "Winds and the orientation of a coastal plane estuary plume," Geophysical Research Letters,37, L19601.

Xia, M., Xie, L., Pietrafesa, L.J. (2007). "Modeling of the Cape Fear River estuary plume," Estuaries and Coasts, 30(4), 698-709.

Dear Reviewer 2,

Thank you for your comments. This manuscript deals with a much more basic problem than assumed by the reviewer. It diagnoses numerical diffusion in coastal models, showing that it is related to advection schemes rather than to discretization or mesh type. There are many other important factors that influence plume propagation in realistic configurations, as explored in the papers cited by the reviewer, but all they work on the top of model numerical factors. We surely agree that the questions raised in the papers cited by the reviewer are important in real-world applications, that many important aspects of plume dynamics were carefully explored in the papers cited by the reviewer. However, the level of numerical diffusion is also important, which is sometimes not fully appreciated. The present manuscript is addressing this question using a specially designed plume configuration that does not characterize a particular river.

It is important to note that all simulations considered in the manuscript

1. do not consider the wind, wave or tidal forcing as all of the listed by Reviewer 2 articles;
2. are done without physical eddy diffusivity (except one to see the difference) to be able to trace the level of numerical mixing accurately;
3. are performed by different models using different schemes.

So, we present a novel test case and diagnostic metrics by which the numerical mixing in ocean models can be quantified. Such a test case can be rather helpful for model developers, especially as it allows comparing models with very different numerical discretizations (in the present paper we consider finite volume C and quasi-B grid models on unstructured and structured meshes, as well as a discontinuous Galerkin finite element model on an unstructured mesh). We believe that the novel test case and metrics are of interest to the model development community.

The references proposed by the reviewer are all examples of excellent work, but they go beyond the present study as they deal with processes not considered by us (e.g. impact of winds, river discharge, or tides) or focus on model assessment in realistic river plume simulations. They do not focus on the systematic assessment of numerical mixing as the present work does. We also note that there are numerous regional studies (some of them are cited in the manuscript), which focus on plume behaviour under different conditions in the different areas worldwide, and we are citing a few of them, including the work proposed by this reviewer.

We have revised the manuscript substantially, shortening and making it easier to read.

I don't think authors have read these references. Clearly Xia 2007 discussed how the numerical schemes impact the plume dynamics, and this submission has overlap with other references. Ideal experiments is very simple, and plume should be investigated with realistic as well. Under the strong river runoff, most mixing scheme won't work well

Dear Reviewer 2,

Suggested reference to Xia et al. 2007 "Modeling of the Cape Fear River estuary plume" does not analyze performance of different tracer advection schemes or limiters with respect to numerical mixing. The individual and coupled effects of the astronomical tides, river discharge, and atmospheric winds were considered to investigate the Cape Fear River Estuary dynamics. On page 699 (right side) the paper provides a brief description of the used EFDC model. Only here the advection scheme is mentioned : 'The model includes the anti-diffusion upwind advection scheme that is more suitable for the plume study than the upwind scheme or the central difference scheme (Berdeal et al. 2002). ’

It looks like a misunderstanding, because we do not consider different mixing schemes, indeed physical eddy diffusivity is set to 0, eddy viscosity is calculated based on k-eps style turbulence closure or set to background value. We consider the numerical mixing level attributed to the different tracer advection schemes and limiters. This is clearly written through the paper.

As we have mentioned in our first reply, we propose the test case and diagnostic metrics by which the numerical mixing in ocean models can be quantified. Spurious numerical mixing in circulation models can destroy stratification and frontal features, and significantly alter the plume dynamics. While considering the effect of physical forcings on the plume dynamics is certainly relevant, it is out of the scope of the present article. The community needs benchmark test cases (see attached Lemarié et al., 2019 summary) that are reproducible, can be compared against analytical solutions and offer the analysis of "isolated" effects (not blurred by the interplay of many different processes as in complex realistic scenarios) and the direct connection to specific numerical choices in the model core. The suggested idealized plume scenario with a unique set of parameters is reproduced differently by different, but commonly used, advection schemes+limiters. And this is not surprising because the plume dynamics in some zones can be characterized by nonlinear flow regimes with sharp frontal boundaries. Simplicity or non-simplicity of idealized experiments depend on the accuracy level you chose as acceptable.

We have read all the papers suggested by the Reviewer. However, their scopes are beyond that of our study. Please, find below a brief report on the suggested articles:

1.Niu, Q., Xia, M. (2021) “The behavior and wind-driven dispersions of two dynamically distinctive limnetic river plumes in a semi-enclosed basin,” Estuarine, Coastal and Shelf Sciences.(In press)

We did not find an article which exactly matches the title, perhaps the Reviewer meant ‘The behaviors of two limnetic river plumes discharging into the semi-enclosed western basin of Lake Erie during ice-free seasons’ .

The article is about wind-driven dynamics of the Detroit and Maumee River sediment plumes in the semi-enclosed western basin of Lake Erie on several temporal scales. In our case study the wind- driven dynamics has not been considered.

2. Niu, Q., Xia, M., Ludsin, S.A., Chu, P.Y., Mason, D.M., Rutherford, E.S. (2018). “High turbidity events in Western Lake Erie during ice-free cycles: Contributions of river-loaded vs. resuspended sediments,” Limnology and Oceanography,00, 1-18.

The article investigates the contributions of river loading (Detroit and Maumee Rivers) versus resuspension to high-turbidity events in Western Lake Erie during ice-free conditions in 2002–2012 using a wave-current forced sediment model (FVCOM based). The major result is that suspended sediment dynamics and high turbidity events in the area were dominated by wind and waves in the offshore regions, and were driven by river loadings near the mouths.

We agree that it is a very important regional study, however, it has focus on sediment dynamics and is hardly relevant to our idealised plume scenario and its major aim.

3. Jiang, L., & Xia, M. (2016). “Dynamics of the Chesapeake Bay outflow plume: Realistic plume simulations and its seasonal, interannual variability,” Journal of Geophysical Research: Oceans, 121, 1424-1445.

The article identifies five types of real-time plume behavior regulated by wind and river discharge. Also it contains some sensitivity experiments related to the grid cell sizes considering fine and coarse grids. The article gives very valuable insights about Chesapeake Bay outflow plume behaviour. However, these five types are defined based on presenting physical conditions (preliminary by wind conditions) and there is no established connection to the numerical scheme performance. Therefore, the topic and analysis of the paper are beyond the topic and aims of our manuscript.

4. “The ideal response of a Gulf of Mexico estuary plume to wind forcing: Its connection with salt flux and a Lagrangian view,” Journal of Geophysical Research, 116, C08035.

The questions posted by the article are:

1) How does wind forcing affect bay water as it encounters the Gulf?

2) How do plume distribution, fluxes, and particle transport change with changing wind conditions?

The questions are important to understand the regional dynamics. However, as they deal with physical forcings, the topic is beyond the current study.

5.Xia, M., Xie, L., Pietrafesa, L.J. (2010). “Winds and the orientation of a coastal plane estuary plume,” Geophysical Research Letters,37, L19601.

The suggested article deals with the Cape Fear River Estuary. and its river plume behavior (type) under different wind forcing and river discharge conditions. Results showed that wind direction, wind speed, and to a lesser extent river discharge contribute to plume transitions from one type to another among six defined major types. This topic is interesting, but not relevant for our study.

6. Xia, M., Xie, L., Pietrafesa, L.J. (2007). “Modeling of the Cape Fear River estuary plume,” Estuaries and Coasts, 30(4), 698-709.

Please, see the comment above.