Review of gmd-2023-236

General comments

Wang et al. describe a new version of the SCHISM hydrodynamic model. They have implemented a total-variation-diminishing (TVD) scheme for ice concentration and tracers, and they have coupled SCHISM to a multi-category column package, Icepack, from the CICE model. They show that the TVD scheme meets model requirements (accuracy, conservation, monotonicity, and efficiency) and generally performs well. They also present results from a coupled multiyear ice–ocean simulation of the Arctic region, showing good agreement with observations.

My main issue is that the paper seems to have two distinct aims—analyzing the TVD scheme and validating the overall model—without doing either in an optimal way. The TVD analysis includes a detailed comparison to upwind and centered schemes, with results that will be unsurprising to anyone familiar with transport schemes. It would be better to compare the TVD scheme to a more sophisticated scheme or schemes (e.g., another second-order monotone scheme), showing that it is either more accurate and better behaved numerically, or able to give similar results at lower computational cost.

The model validation for the Arctic is confined to Section 3.2.2. The results seem promising, but the analysis is short and is limited to ice extent. There is no analysis of the ice thickness simulation, nor is there a comparison to previous SCHISM versions without the TVD scheme and multi-category physics. This makes it hard to assess what has been achieved with the new version.

The paper could be much stronger if it were reorganized, with some material dropped and new material added. For the TVD analysis, most of Section 3.1 could be cut. For the overall model validation, Section 3.2.2 could be expanded. The Introduction could do a better job of motivating the model upgrades, and the Conclusion could give a more complete summary of what has improved and what work remains for the future.

Specific comments and corrections follow.

Specific comments

l. 15 “A more advanced sea ice transport scheme is needed.” The authors give no evidence for this. Rather, the need seems to be for a conservative, monotonic, efficient transport scheme for SCHISM in particular.

l. 18 “Compared with the upwind scheme and a central difference scheme.” This seems like a straw-man comparison; it is not at all surprising that a TVD scheme would outperform these two schemes. More on this below.

l. 35 For sea ice ridging processes in CICE, I suggest citing Lipscomb et al. (2007, JGR) rather than Hunke (2010). Also at l. 111.

ll. 35ff The list of models using CICE or Icepack seems secondary to the main point, and a bit random. It is unclear if or how these various models (e.g., UG-CICE and FESOM2) are related to SCHISM. I suggest first describing SCHISM, the kinds of problems it is used for, the previous implementation of sea ice in SCHISM, and the science goals that explain the need for new and improved components.

l. 59 "sea ice coupled models." Does coupling refer to GCMs and ESMs, or just coupling to ocean models?

l. 61 This is the first use of the term "monotonic" in the main text. Here I suggest defining monotonicity in the context of sea ice transport. Typically, this term refers to schemes that don't introduce spurious new maxes or mins in tracers such as ice thickness or enthalpy. For the case of ice concentration, it would refer to schemes that allow new maxes or mins only when the velocity field is convergent or divergent.

l. 63 "Lipscomb et al. (2004)." Should be "Lipscomb and Hunke (2004)"

l. 67 I don't understand the claim that incremental remapping (IR) is inefficient for unstructured grids. The geometric part of the IR computation scales linearly with the number of grid cells, and the tracer-reconstruction part scales superlinearly with the number of tracers. For CICE and MPAS-Seaice users, the cost of transport is typically not greater than the cost of EVP dynamics and Icepack column physics. If the SCHISM developers opted for TVD in favor of IR, it likely wasn't for reasons of computational efficiency alone. Perhaps they wanted a scheme that was easier to code?

p. 76 "It is unclear…". I don't know why it would be unclear whether or not a scheme is monotonic.

l. 80 For many transport schemes (IR is an exception), the cost increases linearly with the number of variables. Likewise, there is always some cost (usually justified) to imposing strict monotonicity. I'm not sure why these were reasons to rule out FEM-FCT.

l. 81 Here the authors describe the simplicity of the previous version of SCHISM: upwind transport, 0-layer thermodynamics, etc. They could expand on this discussion to say why there was a need for Icepack and other upgrades.

l. 89 "The performance of the multi-class sea ice formulation has not been tested before." This suggests the value of a more complete validation as suggested above.

l. 92 This is where TVD is introduced as a scheme with the desired properties. Can the authors define the method and say when and by whom it was introduced? Does it have a prior history in sea ice modeling?

l. 106 The ITD implementation in Icepack is based on Lipscomb (2001), not Bitz et al. (2001).

l. 120 What is meant by “hydrodynamic core”? Is this the ocean model, or is it something more general than an ocean model?

l. 137 I’m not sure it’s accurate to say that transport is “the main challenge.” Thermodynamics and ridging are challenging too.

p. 141 “a strictly monotone scheme is still desirable”. See the l. 61 comment; it would be better to define and discuss monotonicity earlier.

l. 143 It would be better to introduce the SCHISM model and grid earlier. Please say what is meant by an Arakawa CD-grid. Does the first use of “SCHISM” on l. 144 refer to the SCHISM lake/ocean component?

l. 146 “centroids”. Meaning centroids of triangles, as opposed to centroids of hexagons?

l. 163 It might be helpful to give the reader some examples of Eqs. (7) and (8) in action, showing how they work to preserve monotonicity. For example, one could consider the three cases of (phi_C – phi_U*) = (phi_D – phi_C), 0, and -(phi_D – phi_C), which would imply psi_i = 1, 0, and 0, respectively.

l. 169 “gradient of the central node grad(phi_C)”. Does this mean the quantity grad(phi), evaluated at node C? How is the gradient evaluated? E.g., with a line integral around the adjacent nodes?

l. 181 I’m not sure phi_U* is meant here, since it’s not an edge tracer value. Should this be phi_i, as computed in Eq. 6?

l. 181 The term “sea ice fluxes” is ambiguous. Does this mean fluxes of ice area?

l. 182 Is van Leer limiting applied to *h* and *q*? I think it must be, if *h* and *q* are to be advected monotonically.

l. 205 “Since the thermodynamic part…”. Icepack is a significant step forward for SCHISM, so it would be interesting to know if it improves results compared to earlier model versions, for either the Great Lakes or the Arctic.

l. 211. A time step of 1 s seems unnecessarily short given the size of the triangles (200 m on a side) and speed of the flow (1 m/s).

l. 213 It is predictable that TVD will outperform upwind and centered difference schemes in exactly the ways described. There is no need to include conservation as a metric, since TVD (like upwind) is conservative by construction, whereas centered is not (given that over- and undershoots are clipped). Thus, the following three sections (on accuracy, conservation, and monotonicity) are longer than necessary and not very illuminating.

A more relevant analysis would be to compare TVD to incremental remapping (if the authors were able to set up similar test problems in CICE or MPAS-Seaice) or another second-order monotone scheme. In the case of IR, it could be interesting to show that TVD gives similar results at lower cost.

l. 289 The high-resolution Great Lakes simulation is a good problem for comparing TVD and upwind. The results shown in Fig. 6 are quite convincing. For this reason, I think the authors could leave out the simple problems in Section 3.1 and let Section 3.2.1 make the case for TVD over upwind.

It is interesting that the TVD method reduces the overall model cost (compared to upwind) by limiting diffusion of ice area. How much time is spent in the transport solver alone for each of the two transport schemes, and how does this compare to the total model time?

l. 304 I suggest "compare" instead of "qualitatively compare", since the comparison is not merely qualitative.

l. 315 Since this section focuses on general model validation (rather than a validation of TVD), I suggest expanding it and making it an entire section rather than a subsection. Also, it would be useful to see how the new model version compares with the older, simpler version.

l. 318 Is the sea ice time step just 100 s? This seems unnecessarily short if the minimum grid cell size is 6 km, assuming a max speed of ~1 m/s.

l. 328 I am not sure what is meant by "the generic length-scale equation as k-kl."

l. 330 How is TVD^2 related to the TVD scheme implemented for the sea ice model?

l. 341 "which may be influenced by the initial conditions as we did not get all tracers, such as sea ice salinity and enthalpy, from HYCOM." I can think of many reasons why the first peak might not line up with the observed value. I'm not sure why initial tracer values are singled out as an explanation.

l. 350 Why was FESOM2, as opposed to some other model, chosen as a standard for comparison? How similar was the FESOM2 configuration?

l. 354 “the simulated sea ice extent often increases faster in autumn than observation.” This isn’t obvious from Fig. 7a. I just see one year (1994) when the modeled September min is significantly greater than observed.

l. 361 Typically when comparing two models, one would force them over the same integration period. If FESOM runs are available from 1994–1999, I would suggest using those. If not, then it might be better to leave out the comparison.

l. 368 It’s helpful to see these spatial patterns of sea ice concentration biases. Would it be possible also to show plots of sea ice thickness compared to observations?

l. 371 Do the authors know why the ice edge is too far advanced on the Atlantic side, and not far enough on the Pacific side? Is this likely an ocean model bias?

l. 381 The melt pond hypothesis is interesting. Is it possible to test this idea by, for instance, turning off melt ponds or using different precipitation forcing?

l. 393 The discussion section is short and includes some material (e.g., grid choice) that would fit better earlier in the paper. It doesn’t shed new light on the Section 3 results. I would suggest leaving it out.

l. 405 I doubt that “remarkable” is the right word here. Again, I don’t think the centered difference scheme adds value to the Section 3 analysis.

l. 410 This is an odd place to introduce the Casulli et al. scheme. Maybe do this earlier, in Section 2.2 or 3.1.

l. 417 The conclusion is short and cursory. It would be better to include a discussion of how the addition of Icepack and the TVD scheme have improved SCHISM compared to the previous model version.

l. 460 The reference list is incomplete and contains some errors. For instance, there is no Gurvan et al. (2022) or Campin et al. (2023).

Minor corrections

l. 23 “the satellite” -> “satellites”
l. 25 “dramatically” -> “dramatic”, “Sea ice” -> “sea ice”
l. 29 “the sea ice models” -> “sea ice models”
l. 87 “The Great Lake” -> “the Great Lake”

l. 109 (and elsewhere): “traces” -> “tracers”
l. 132 (and elsewhere): “Where” -> “where”
l. 139 (and elsewhere): Check punctuation with equations. Here, the comma should be a period.
l. 162 “Van-leer” -> “van Leer”. Also l. 197.
l. 183 “as does CICE” -> “as in CICE”
l. 324 “manning” -> “Manning”
l. 325 “The” -> “the”
l. 335 “Sea Ice Concentrations” -> “sea ice concentration”

This is not a complete list. There are many minor typographical and grammatical errors that should be cleaned up in the next version.