RC2

General comments:

This study develops an operational forecast system with a three-dimensional lake hydrodynamic model and validated the performance of hindcast experiments in Lake Erie. To conduct a robust forecast, the study separated 24-h and 240-h forecast simulations, and update the restart file for 240-h forecast every day. The data retrieval, numerical simulation, and validation are automated. This study is a development of a forecast system rather than a new model, but it is still of importance from the perspective of research implementation to society. However, some information for long-term steady operation is lacking, and providing us with such information would be helpful to other models/system developers, which are listed in the following major comments.

> **Reply 1:** Dear Dr. Tokuda, thank you for taking the time to review our manuscript, and your comments are very supportive of our work. In terms of long-term steady operation, if you are indicating the system stability in couples of months, we have added a comparison of stitched 24-h model forecast and continuous model run with 2 months of retrospective meteorological forecast in Appendix A (Fig. A1). So far, the system has been stably running under supervision over 2020-2021 without any runtime errors in the model and keps generating iterative forecast results every day. We have added the information about the supervisor of COASLTINES in lines 207-208,
>
> "The authors (supervisors of COASTLINES) and Queen's ITS monitor forecast results and maintain system operation."

Specific comments:

How did you set the initial value of the model? You mentioned that the model was 'cold started' on day 99 (line 116) and generates a restart file with a 24-h forecast simulation (line 185), but some other sentences imply that the model was initiated on another date (e.g., lines 288 and 299). I imagine that "initiated" means the model was just restarted in the consistent simulation (if so, only lead (forecast) time information is enough), but the authors are requested to clearly explain the system setup because such information is useful for other forecast systems without data assimilation.

Reply 2: Thank you for the correction and sorry for the confusion here. Your understanding is correct. To avoid the misunderstanding, we replaced ‘initiated’ with ‘hot-started’ in the revision (e.g., captions of Fig. 5, 8, and 10). We also modified the content in 2.5 System operation to clarify the approach (lines 188-191, lines 193-195):

“The model advances every day according to the 24-h forecast simulation and terminates by generating ‘re-start’ files. These files are then used to hot-start the 240-h forecast simulation and the 24-h simulations for the next day. The input files for the 240-h forecast simulations are iteratively replaced by the new 240-h meteorological forecast generated each day.”

“The long-term stability of employing daily ‘hot’ restarts can be seen in a comparison between simulated temperature profiles from a continuous run and that from stitching together the 24-h hot-start simulations (Appendix A; Fig. A1).”

What is the difference of the meteorological forcing data between 15km and 25km? If your system can work in time with 15-km data, I think you can focus on the results of 15-km data only because the difference of results seems to be negligible. Or is there any known problem with 15-km data?

Reply 3: 15 km and 25 km are the different horizontal resolutions of the meteorological forecasts. We wanted to test the sensitivity of the model to meteorological input resolution, but it turns out that the discrepancies in 15 km and 25 km meteorological forecast are minor, and for the large lake (e.g., Lake Erie), the forecast model is not sensitive to these minor discrepancies (Fig. 3, 7). The evaluation metrics (Table 2) show that the forecast outputs from 15 km model did not show lower deviation as expected, and sometimes even had higher bias. However, the 15 km model was more sensitive in predicting the upwelling event (Fig. 8, 9, D2). Thus, we suggest meteorological forecast with 15 km should be implemented to detect the meso-scale phenomena like upwelling. We have added the discussion in lines 452-454.

“The meteorological forecast from the 15 km and 25 km GDPS models did not show discrepancies (Fig. C2-5) and the evaluation metrics indicate that forecast results were largely insensitive to the meteorological inputs in Lake Erie (Fig. 3, 7). However, the 15 km model better predicted the mesoscale upwelling event (Fig. 8, 9, D2).”

You showed one-year results, but I have an interest in the long-term stability of the system operation. Do you reset the initial condition every year?

> **Reply 4:** We did not reset the initial condition every year. So far, we have not seen any runtime error in the model for 2020-2021 winter and spring. Moreover, we have run continuous hindcasts with AEM3D over 2002-2012 and the model does not show significant drift.

2 Data and methods

2.2 Model description

Line 96: Which programming language is the AEM3D written in? You mentioned that the wrapper code is written in Python, but it has no advantage on computational efficiency. In addition, you pointed out that previous hydrodynamic models are difficult to apply to the hind- and forecast applications due to the computational cost in line 42, how did you solve the problem?

> **Reply 5:** AEM3D is written in Fortran which has advantage on computational efficiency and the model itself can complete the short-term forecast (up to 10 days) at a reasonable real/run time ratio (0.5h for 24h simulation, and 4h for 240h simulation). However, we treat the hydrodynamic driver as a black box. We agree that Python is not computationally efficient, but we also pointed out that it is the complex and time-consuming setup and calibration procedure in the hindcast application of the model that limits the implementation of models for hydrodynamic forecasts (lines 42-47).
>
> "In the case of hindcast applications, the complex and time-consuming setup and calibration procedure, of these models, can result in a significant time lag (months to years) between when a project is initiated and when the model results are communicated to stakeholders. This delay severely limits the utility of computational models for policy and management decision making. For better application of these powerful computational tools, the ability for rapid monitoring and simulation forecasts should be established."

The manuscript underlines the importance of the automatic workflow developed in this study in the Introduction, which could accelerate the data acquisition and model setup procedure, and eliminate the hurdle in generate iterative forecast results (lines 69-72).

"In the present study we developed and tested the COASTLINES (Canadian cOASTal and Lake forecastINg modEl System; https://coastlines.engineering.queensu.ca/) lake-model application workflow, that rapidly accesses near real-time online data (weather forecasts, water level and temperature observations) for automated model forcing, execution and validation. "

2.3 Model setup and meteorological forcing variables

Line133: I have three questions.

(1) The mass balance of lakes is described in + Precipitation – Evaporation + Riverine inflow – Riverine outflow ± Groundwater infiltration/seepage. Did you consider the groundwater component? If not, explicitly describe that assumption.

Reply 6: No, we did not consider the groundwater component. Due to the large amount of water stored in Lake Erie and its large lake surface area, the hydrodynamic simulation of the lake usually does not include the groundwater component since it plays a minor part in the water circulation.

(2) In addition, is Precipitation – Evaporation balanced in Lake Erie? This budget controls the seasonal variability in water level, but mass imbalance may cause a problem in longterm operation.

Reply 7: Yes, we agree that the water balance assumption may affect the model results in the long-term forecast. But in terms of the short time scale (10-day forecast) the oscillation of water level was mainly caused by surface seiches (Trebitz, 2006). The Precipitation – evaporation balance does not affect the water level significantly in such a short time scale. Unfortunately, there is no way to reset the water level within AEM3D using the restart file; therefore, in our projections we adjusted the predicted water level according to the real-time observed gauge levels. Future work will employ machine learning to forecast inflows/outflows so as to achieve a water balance.

We added this information into lines 136-139, "In this pilot application, the Lake Erie inflows and outflows, which roughly balance, are neglected, however evaporation and precipitation are accounted for in the water balance. Over short timescales (<10 days), the contributions from evaporation and precipitation to water level change are minor, with water level oscillations resulting from storm surges and surface seiches (Trebitz, 2006)."

(3) Even if the riverine in/outflows are balanced, can you ignore the effect on the fluid velocity field near the inlets and outlets? If you can, add the reference.

Reply 8: Yes. We ignored the effect on the fluid velocity filed near the river months. Because Lake Erie has such a large scale, this is a very small part of the model domain and we did not see this simplification to cause any errors or affect the hydrodynamic predictions, except the water level forecast near Bar point (lines ??).

3 Results

You showed the confidence shade in the figures of time series, how did you do conduct ensemble simulations? This question is related to the reliability of the system if it is operated as a warning system to society.

Reply 9: We did not conduct the ensemble simulations, but using the averaged statistical metrics gained from the previous forecast evaluation to indicate the confidence level. In the beginning of the Results, we have modified the text in lines 219-223 as

"The water level statistical metrics (RMSD and RE) were ensembled and averaged over April to September 2020. The 24-h and the 240-h forecast lake surface temperature and temperature profiles, from the models, were also validated against real-time lake buoy data and daily averaged satellite imagery. The timeseries and spatial MBD and RMSD (t-RMSD, t-MBD and s-RMSD, s-MBD) were ensembled and averaged over July to September 2020."

3.2.1 Lake surface temperature

Line 273: Those results are interesting; longer forecast time does not increase the error for surface temperature (thermodynamics) even with consistent bias according to Fig. 6 but does the error for water level (hydrodynamics) according to Fig. 4 and 5. Can you discuss the difference? Comparison between the model bias and forecast time would be helpful in this respect.

Reply 10: Thank you for pointing this out. The comparisons between model bias and forecast time have been shown in Fig. 3 and 7 for water level and lake surface temperature, respectively. The water level oscillations in Lake Erie are mostly due to the surface seiches (Trebitz, 2006), which depend largely on the wind, compared to the lake temperature, which depend mainly on the air temperature. The wind forecast has larger uncertainty as the forecast time extends (Fig. C4, 5); thus, we can see the obvious growth of bias against forecast time in water level (Fig. 3) but not in lake surface temperature (Fig. 7). We have supplied some discussion on this topic (lines 450-452 and lines 460-461).

"The rapid response of the water level to windstorms (Hamblin, 1987) could result in the effects of aliasing and forecast error being passed to the water level, leading to the growth of RE against forecast time (Fig. 3)."

"The growing bias in air temperature, with forecast time, does not affect the lake surface temperature (Fig. 7), presumably owing to the buffer effect of surface mixing layer (Schertzer et al., 1987). "

Line 280: Why could you conclude the underestimation is due to ignoring river inputs? The underestimation occurs on the east side of the inlet from the Detroit River. Can you have a consistent discussion between line 223 and here?

Reply 11: To make a consistent discussion, we have moved the discussion about bias induced by neglecting inflows into 4.3 (lines 462-469).

"Neglecting the inflows and outflows in the predictive simulation could induce bias in the forecast. The overestimation of water level fluctuation range near Bar Point (Fig. 4f) may result from neglecting the large Detroit River inflow, which regulates the seiche magnitude. The inflows also adjust more rapidly to air temperatures compared to deep lake waters. Thus, the up to 2 ℃ cold bias in coastal regions of the western basin (Fig. 8 m-t, Fig. D2) could be induced by neglecting the heated flux from two major inflows (i.e., Detroit River and Maumee River) of Lake Erie."

Fig. 8: What is the main reason for the consistent underestimation in lake surface temperature?

Reply 12: It is induced by the bias in air temperature forecast as the forecast lead time increase (e.g., Fig. C5). We have added the discussion in lines 456-458, "The 168-h forecast meteorological data overestimated wind speeds by up to 10 m s^{-1} (Fig. C4), and bias in the air temperature forecast (Fig. C5) may cause the consistent warm bias (up to 3℃) in forecast lake surface temperature (Fig. 8)."

4 Discussion

4.1 Bias and uncertainty

If the system developed in this study focuses on the forecast of some critical events like coastal up-welling and storm surge as discussed in the Sects. 4.1 and 4.2, could you move this Sect. 4.1 after the Sect. 4.3? The current Sect. 4.1 discussed the mean RMSD and compared it with a previous study incorporating data assimilation, but the data assimilation corrects only the initial condition of state variables in a model; not boundary conditions including meteorological forcing data. On the other hand, some of the critical events are caused by extreme atmospheric conditions in my understanding. So, can you discuss the model uncertainty and further improvement separating into initial-value and boundary-value problem?

Reply 13: Thank you for the suggestion. We have moved Sect. 4.1 to the last part of Discussion (4.3) and separate the discussion into initial condition-induced errors and boundary condition-induced errors (lines 429-461).

Technical corrections:

2.3 Model setup and meteorological forcing variables

Line 116: How did you set the initial condition for the water level?

Reply 14: AEM3D does not allow for adjustment of the water level in the restart file. Therefore, we project the changes in water level at the gauges, using the observed gauge value adjusted by the predicted water level change according to the model. This is done in post processing (within the automatic workflow). We have added the explanation in lines 195-198. Eventually the model will be further developed to compute a proper water balance.

"At present, the initial water level cannot be modified using the AEM3D re-start files. Therefore, to account for long term drift in surface water level, we used real time gauge observations as the datum point for water level forecasts (automatically performed by MATLAB in post processing) and only consider errors resulting from simulation of storm surges and seiches, as opposed to those from seasonal changes in mean lake level."

Line 119: What is "which is CFL = (Hodges et al., 2000)?

Reply 15: We are sorry about the typo here. It should be CFL = $\sqrt{2}$ (line 121).

Line 121: "and net longwave radiation" (the former one) -> "and downward longwave radiation"? Because net (downward -upward) longwave radiation is calculated within the model as you mentioned.

Reply 16: We have corrected the text here as you suggested.

3.2.1 Lake surface temperature

Line 252: Lake "S"urface temperature -> Lake "s"urface temperature

Reply 17: We have corrected the text here.

Line 288: needs a punctuation after "day 251".

Reply 18: We have added a punctuation here. Thank you for correction.

4.2 Prediction of coastal up-welling for fishery and drinking water management

Fig. 11:

(1) Correct the caption of the colorbar (remains selected?), and the same problem happened in Fig. D1.

Reply 19: We have corrected both figures.

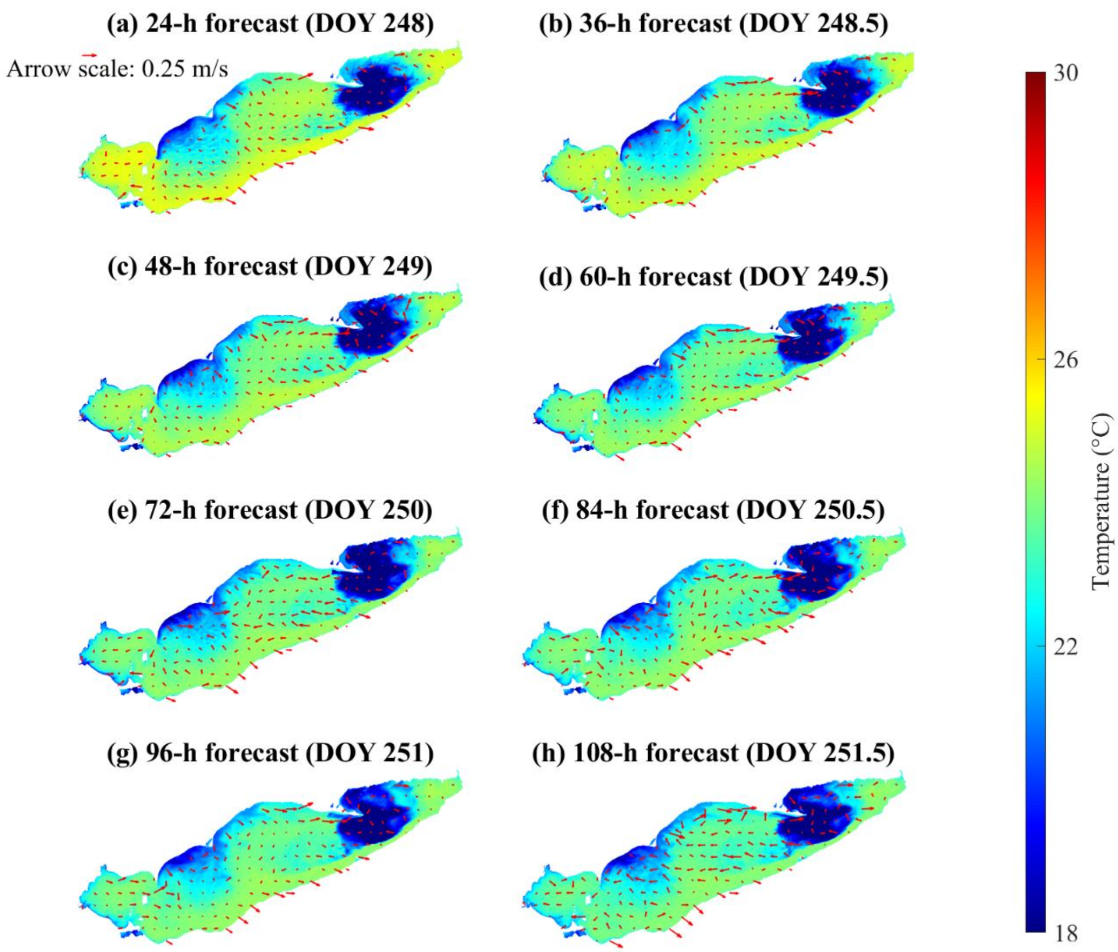

Fig. 11 Color maps showing the forecast depth-averaged temperature throughout the lake. The red arrows represent forecast depth-averaged currents. The model results are from the 240-h forecast model hot-started on day 247.

(2) Can you show the observation data?

Reply 20: Because this figure presented as a depth-averaged temperature, there is no observation data to compare against.

4.3 Prediction of storm surge events for public safety

Fig. 12:

(1) Can you show the spatial distribution of the 24-h and 96-h forecast to compare with the time series in (d)?

Reply 21: As you suggested, we have added the map of water level change from 24-h forecast and 96-h forecast in the Appendix E.

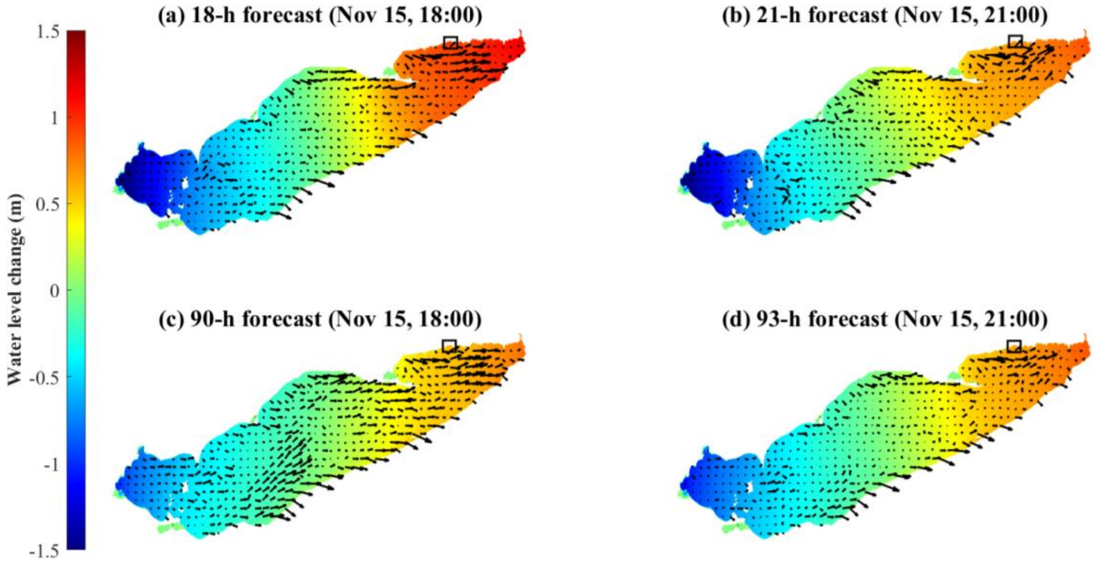

Fig. E1 Spatial distribution of water level change from forecasts hot-started on Nov 15th (a, b) and Nov 12th (c, d). The reference water level is the observation at Nov 15th 00:00. The black arrows are depth-averaged mean current fields. The black squares in the upper right corners of each map indicate the location of Port Dover (Fig. 12d).

(2) According to (a) to (c), the forecasted water level is highly dependent on forecast time, and can you show the relationship between water level and forecast time? (It does not seem to be saturated even if the latest data is used according to (a)

Reply 22: Yes. Due to the windstorms, the water level in Lake Erie is highly dependent on surface seiche, with period ~ 14 h (Trebitz, 2006;Mortimer, 1987). Panels (a) to (c) show the forecasted basin-scale surface seiche during a windstorm. The relationship between water level and forecast time was shown at one location (Port Dover), where we have footage showing the effect of flooding, in panel (d). Could you please specify what kind of information you want us to visualize?

Reference

Hamblin, P. F.: Meteorological forecing and water level fluctuations on Lake Erie, J. Great Lakes Res., 13, 436-453, 10.1016/S0380-1330(87)71665-7, 1987.
Mortimer, C. H.: Fifty Years of Physical Investigations and Related Limnological Studies on Lake Erie, 1928–1977, Journal of Great Lakes Research, 13, 407-435, https://doi.org/10.1016/S0380-1330(87)71664-5, 1987.
Schertzer, W. M., Saylor, J. H., Boyce, F. M., Robertson, D. G., and Rosa, F.: Seasonal Thermal Cycle of Lake Erie, Journal of Great Lakes Research, 13, 468-486, https://doi.org/10.1016/S0380-1330(87)71667-0, 1987.
Trebitz, A. S.: Characterizing seiche and tide-driven daily water level fluctuations affecting coastal ecosystems of the Great Lakes, J. Great Lakes Res., 32, 102-116, 10.3394/0380-1330(2006)32[102:CSATDW]2.0.CO;2, 2006.