

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2021-34-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-34

Anonymous Referee #1

Referee comment on "An automatic lake-model application using near-real-time data forcing: development of an operational forecast workflow (COASTLINES) for Lake Erie" by Shuqi Lin et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-34-RC1>, 2021

Lin et describes and evaluates a forecasting system for predicting 3-D thermal structure of Lake Erie. The manuscript is an evaluation of a true forecasting system (i.e., it is evaluating a set of forecasts of the future rather than mimicking the forecasting process with historical data). It uses forecasted meteorology from Environment Canada Global Deterministic Forecast System to drive the model. The model used (AEM3D) had previously been applied to Lake Erie so the novelty of the paper is using the model with forecasted meteorology. I really appreciated the discussion sections on how the forecasting system could potentially be used to anticipate critical events for decision makers, emergency managers, and users of the lake. I visited the website for the project and it is indeed up-to-date. The paper evaluates the performance of the forecasts for a three-month period of time, although the manuscript discusses forecasts from outside this period.

The authors highlight the automation provided by a Python script but this seems to lack novelty (thought it appears to get the job done). The model is a standardly used model, downloading one meteorology source and converting to another is straightforward, and using a task manager to run a job is routine. The application to Lake Erie is new but the manuscript is introducing a named forecasting system (COASTLINES) with the only model development being a Python script used to execute the model and download observations (actually also Matlab scripts). Further, the code for running COASTLINES is not provided. Overall, the novelty of COASTLINES beyond the application to Lake Erie needs to be better motivated in the introduction.

The manuscript only uses observations to evaluate the model. It does not perform data assimilation as other forecasting system do. The workflow highlights the automation of the evaluation using the observations but does not offer a way that the evaluation is used to improve the model or the forecasts. Therefore, it isn't clear why the automated evaluation is necessary. It would be great to explore how a feedback between the evaluation and the forecasts can be developed.

The forecasts lack a representation of uncertainty. Uncertainty in forecasts is increasingly the state-of-the-art. There is reference to uncertainty in the figures but the manuscript does not describe how uncertainty is estimated. At minimum, the discussion needs to address the lack of uncertainty in the forecast and explore how it might be included in the forecasts.

The argument for why data assimilation is not necessary could be stronger. They argue that it would only potentially decrease the RMSE by ~half (0.7C) by citing another study that used data assimilation. However, how does the reader know whether this is a meaningful magnitude?

Instead of putting the Python script in the Appendix, I recommend putting them in a repository like Zenado. That would allow someone to use the scripts without having to cut and paste from the Appendix. Furthermore, the manuscript highlights the Python code but also has a dependence on MatLab for foundational parts of the workflow (i.e., line 131-132 – the conversation of weather forecasts into the AEM3D input). Therefore, it is a Python + Matlab + Windows Schedule workflow. Additionally, the code that is provided is only for retrieving the observations. The code to run the forecasting system is missing. Overall, I do not think that the availability of code and model output meets GMD's standards.

The term hindcast is used throughout the manuscript but is not defined. It would be help to define exactly what a hindcast is.

Line 440 states "To facilitate further development of open-access predictive modelling systems", which is a major oversell of COASTLINES as being in the group of open-access predictive modeling. The hydrodynamic model requires a license and the code to run the forecasting system is not made available.

Can the authors point to a manuscript that demonstrates that the re-start works or provide a figure that shows that a restarted run matches a run that was continuous (i.e., same length of simulation but without stopping and restarting)? Some hydrodynamic models are designed to be run from a cold-start and have internal variables that are not saved – thus preventing a true restart.

The manuscript highlights the use of the Environment Canada Global Deterministic Forecast System (GDPS) product. Could other freely available forecast model output be used? What about NOAA's Global Forecasting System or NOAA's Global Ensemble Forecasting system?

Specific comments

Line 47: Are there 1-D water temperature forecasting systems that can also provide context in the introduction and discussion?

Line 63: I recommend starting a new paragraph here

Line 95: Is there a specific version of the AEM3D that was used? Without that the forecasting system is not reproducible.

Line 119: what is CFL =?

Line 156: What happens when there are run-time errors?

Line 184: Is the restart file used to generate a 216-hr forecast since the first 24-hr have already been generated?

Line 189: The Windows Task Scheduler is another dependency of the forecasting system. Can the forecasting system only run in a Windows environment?

Line 224: The sentence refers to the estimation of uncertainty in the model forecast but the manuscript lacks methods describing the uncertainty estimation process.

Line 260: What forecast horizon do these numbers refer to?

Line 329: Readily what? (missing word)

Fig 2: What is Phenomena detection (Supervisor)? It is not mentioned anywhere in the manuscript

Fig 2: The caption says "Daily Python workflow" but the text states that Matlab is also used.

