

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2022-174-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2022-174

Anonymous Referee #1

Referee comment on "Prediction of algal blooms via data-driven machine learning models: an evaluation using data from a well-monitored mesotrophic lake" by Shuqi Lin et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2022-174-RC1>, 2022

Review

Prediction of algal blooms is important for the conservation of natural water systems. The authors proposed a two-step ML approach to overcome the limitation of insufficient observations. However, there are still some problems need to be solved:

- The authors mentioned some machine learning models in the Introduction. These are important in the development of algal bloom prediction. The authors should analyze the disadvantages of these ML models and the improvements of their own model.
- The literature review of ML models is too simple, which makes it difficult to find the development of models.
- Page 3, Line 73. The significance of designing three workflows needs to be further clarified.
- Page 3, Line 80. Why do the authors use GBR and LSTM?
- The advantage of two-step method is accurate prediction when observations are insufficient. However, workflow 1 performs better than workflow 2 or 3 (Table 1). From this comparison, the two-step method is not an important step that affects the accuracy.
- From Fig. 3 (e.g., Kappa scores), the PB model also works well. What is the advantage of ML models?