Comparison of Eulerian and Lagrangian transport models for harmful algal bloom forecasts in Lake Erie

Xing Zhou[1,2], Mark Rowe[3], Qianqian Liu[4], Pengfei Xue*[,1,2,5]

[1]Department of Civil, Environmental and Geospatial Engineering, Michigan Technological University, Houghton, MI 49931; United States

[2]Great Lakes Research Center, Michigan Technological University, Houghton, MI 49931; United States

[3]NOAA Great Lakes Environmental Research Laboratory, Ann Arbor, MI 49931; United States

[4]Department of Physics and Physical Oceanography, University of North Carolina Wilmington, Wilmington, NC 28403; United States

[5]Environmental Science Division, Argonne National Laboratory, Lemont, IL, 60439; United States

Highlights

- Conducting a comprehensive evaluation of Lagrangian, Eulerian, and hybrid transport models for Lake Erie CHAB hindcasts.
- Examining dynamic response of currents to weather-scale wind events and associated CHAB transport in the short-term forecast
- Quantifying the competing influence of vertical turbulent mixing and algal buoyancy on surface CHAB intensity
- Discussing strengths and limitations of the three types of model in ecosystem forecast applications

Abstract

Lake Erie has experienced a re-emergence of cyanobacterial harmful algal blooms (CHABs) since the early 2000s, posing significant socioeconomic and ecological consequences that impact drinking water, human health, fisheries, tourism, and water quality. As predicting CHAB intensity and spatial distribution is critical to Lake Erie ecosystem management, this study focuses on a comprehensive evaluation of Lagrangian and Eulerian transport models for Lake Erie CHAB forecasts, including 1) a Lagrangian particle model (LPM), 2) an Eulerian tracer model (ETM), and 3) a property-carrying particle model (PCPM) that utilizes the hybrid Eulerian-Lagrangian approach. We evaluated the models' performance against the latest high-resolution satellite product from the European Space Agency's Sentinel-3 OLCI sensor over 24- to 240-hour hindcasts for each CHAB occurrence in three consecutive CHAB seasons (2017-2019). We examined the relative contributions of horizontal transport, vertical turbulent mixing, and algal buoyancy on the CHAB inter- and intra-day variability. In the short-term forecast, we emphasize the highly dynamic reaction of currents to weather-scale wind events that are crucial to CHAB transport. While statistical skill assessments show that these three transport models attain comparable levels of hindcast accuracy, we explore the advantages and disadvantages of each model in the context of general biophysical modeling. In particular, the fact that the ETM and PCPM perform as well

as or better than the LPM sets up a promising path to developing more biological realism in future operational forecast models using Eulerian or hybrid approaches.

Key Words: Harmful Algal Bloom; Hydrodynamics; Transport Model; Lake Erie; Great Lakes

1. Introduction

Lake Erie is the most productive lake among the five Laurentian Great Lakes of North America. The re-emergence of cyanobacterial harmful algal blooms (CHABs) since the early 2000s has been known as one of the most severe environmental problems in the Great Lakes (Watson et al., 2016; Liu et al., 2020) as a result of eutrophication conditions caused by the excessive nutrient input primarily from agricultural use (Baker et al., 2014). The most severe CHABs originate in the shallow western basin of Lake Erie (Fig. 1), associated with nutrient loading from the Maumee River in the southwest of the basin (Kane et al., 2014). CHABs in Lake Erie are dominated by *Microcystis aeruginosa*, which produces a family of algal toxins known as microcystins that are dangerous to humans and animals. *Microcystis aeruginosa is* a species of freshwater cyanobacteria that can regulate the buoyancy of its colonies, with most of the colonies being positively buoyant in Lake Erie (Uyl et al., 2021).

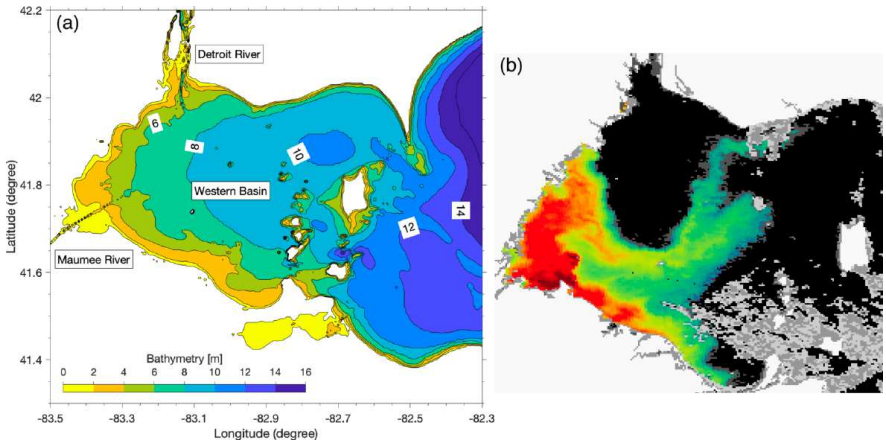

Fig. 1 (a) Bathymetry of the western basin of Lake Erie. (b) A satellite image from the Sentinel-3 OLCI sensor shows an example of a severe CHAB event in the western basin of Lake Erie on August 7th, 2019.

CHABs have high socioeconomic and ecological costs in Lake Erie, impacting drinking water, human health, fisheries, tourism, and water quality. For example, on August 2nd, 2014, the city of Toledo issued a "do-not-drink" notice for their drinking water system that affected 0.5 million people for three days, as a result of contamination of treated drinking water with microcystins, associated with a severe CHAB event (Stumpf et al., 2016; Steffen et al., 2017). The severe impact caused by CHABs promoted the development of an improved short-term forecast system for CHAB distribution and transport (Rowe et al., 2016).

As physical processes play a dominant role in explaining short-term bloom variability (Rowe et al., 2016), short-term CHAB forecasts often employ a mass-conserving, advection-diffusion model. Such a model is typically forced by a hydrodynamic forecast, with CHAB spatial distribution initialized from satellite remote sensing, to predict the bloom location and intensity several days into the future. The partial differential equations governing algae transport can be solved using a Lagrangian approach, an Eulerian approach, or a combination of the two (i.e., a hybrid Eulerian-Lagrangian approach).

In Lake Erie, both the Lagrangian and Eulerian approaches have been tested for CHAB hindcast simulations (Wynne et al., 2013; Rowe et al., 2016; Soontiens et al., 2019). The Lagrangian particle model (LPM) describes motion by following individual particle trajectories as the particles move through space and time. Each particle represents a certain amount of planktonic or chlorophyll mass and is transported by three-dimensional velocity components and turbulent diffusivity output from a hydrodynamic model. The CHAB condition is therefore expressed as an accumulative behavior of phytoplanktonic particles. As each individual particle can contain a set of attributes (e.g., colony size, buoyant velocity) with corresponding physiological traits and behavioral traits, this approach allows for considering the colony variability of CHABs (Henrichs et al., 2015). The LPM is limited in its ability to represent continuous concentration fields because of the need to count particles within a control volume to represent a concentration. Unlike the LPM in which mass is transported as discrete particles, the Eulerian tracer model (ETM) expresses planktonic biomass or chlorophyll as tracer state variables and calculates their concentrations in each model grid cell. The Eulerian specification of a field calculates the local rate of change at fixed locations (grid cells) determined by advection and turbulent diffusion as well as its biophysical source and sink terms.

While a few previous studies have assessed short-term hindcast simulations of Lake Erie CHABs using Lagrangian or Eulerian approaches, drawing an insightful conclusion on model performances across these approaches remains challenging. This is due to the fact that these models were evaluated using different skill measurements, different biophysical settings, different forecast periods, and/or different model versions. For example, Wynne et al. (2013) focused on a two-dimensional (2D) LPM that was initialized with satellite remote sensing data and driven by a hydrodynamic model output to hindcast the CHABs for 2008-2010. Rowe et al. (2016) took a similar approach but applied a 3D LPM for each CHAB event in the 2011 CHAB season for up to 240-hour hindcasts. By taking into account the vertical distribution and buoyancy of *Microcystis* colonies, the 3D version of the LPM delivered more skillful hindcasts than its 2D version. More recently, a 3D ETM was compared to a 2D LPM (rather than a 3D LPM) that limits particles to the lake surface by Soontiens et al. (2019). The comparison revealed that the 3D ETM outperformed the 2D LPM by a wide margin. However, a 3D ETM and 3D LPM were not directly compared. In contrast to the 120-hour and 240-hour hindcasts from Wynne et al. (2013) and Rowe et al. (2016), Soontiens et al. (2019) assessed shorter 48-hour hindcasts from July 27th to October 8th, 2017.

Therefore, this study aims for a comprehensive model skill assessment, directly comparing 3D versions of Eulerian and Lagrangian CHAB forecast approaches, including assessing a hybrid Eulerian-Lagrangian approach (see section 2.4 for the model description). These models were run in the latest 3D version at the time of the writing and were assessed under identical biophysical conditions and statistical skill metrics. For each CHAB occurrence in three consecutive CHAB

seasons (2017-2019), we evaluated model performance against the latest high-resolution satellite product from the European Space Agency's Sentinel-3 OLCI sensor over 24- to 240-hour hindcasts. Then, we established a statistical skill assessment based on all 43 CHAB events to identify each model's strengths and limitations.

2. Methods

2.1 Hydrodynamic model

The Finite Volume Community Ocean Model (FVCOM) is a three-dimensional, free-surface, primitive-equation hydrodynamic model that solves the integral form of the governing equations on an unstructured, sigma-coordinate mesh. The advantage of an unstructured grid mesh for shoreline fitting and local mesh refinement makes the model particularly popular in applications to coastal waters (Xue et al., 2014, 2020, 2022; Ibrahim et al., 2020; Huang et al., 2021a, b). FVCOM has been applied in many coastal systems characterized by geometric complexities and highly variable flow patterns, including various applications to Lake Erie (Kelly et al., 2018; Rowe et al., 2019; Ye et al., 2020).

The Lake Erie (LE)-FVCOM has an unstructured grid mesh of 6106 nodes and 11509 elements, with the water column divided into 20 uniform-thickness terrain-following sigma layers (Fig. 2a). The mesh has a grid resolution of ~2.5 km in the central basin, ~1.5 km in the western basin, and ~0.5 km in Maumee Bay and the islands. The open boundary conditions include primary inflow from the Detroit River and outflow through the Niagara River with specified hourly water levels using the NOAA/NOS gauges at Gibraltar (9044020) and Buffalo (9063020). The LE-FVCOM is driven by hourly atmospheric forcing from the High-Resolution Rapid Refresh (HRRR) model, a cloud-resolving and convection-allowing weather forecast and data assimilation system running in real-time at a 3-km grid resolution (Benjamin et al., 2016).

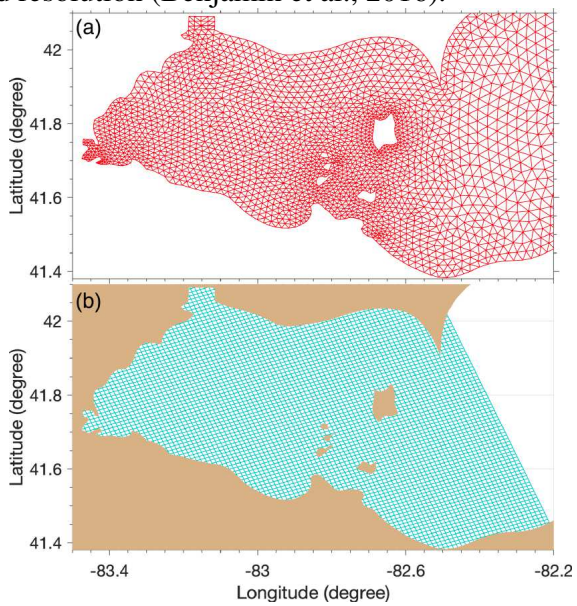

Fig. 2. (a) An enlarged view of the western portion of the triangular mesh domain used by the FVCOM, Lagrangian, and Eulerian tracer models. (b) A view of the PCPM grid with 1 km resolution.

2.2 Lagrangian particle model (LPM)

The LPM described by Rowe et al. (2016) is used for the Lake Erie Harmful Algal Bloom Forecast system (https://coastalscience.noaa.gov/research/stressor-impacts-mitigation/hab-forecasts/lake-erie/, accessed August 2nd, 2021), which was originally described by Huret et al. (2007) and Churchill et al. (2011). The experimental version was sometimes referred to as the CHAB Tracker. The same LPM was used to hindcast the probability of microcystin concentration exceeding public health advisory guidelines (Liu et al., 2020). In the LPM, CHAB extent and intensity (described as cyanobacterial chlorophyll concentration) are represented using Lagrangian particles. Each particle represents a specific chlorophyll mass ($10^{10}\ \mu g/\quad cle$) and chlorophyll concentration in each grid is determined by counting the number of particles in each FVCOM tracer control element. The LPM calculates concentration using the same horizontal mesh as FVCOM but using constant-thickness (1 m) z-layers instead of sigma-layers. Advection of particles is governed by the following equation:

$$\frac{dX(t)}{dt} = V(X(\ ),\ ) \tag{1}$$

Where $X(t)$ is the three-dimensional particle position at time t, $V(X(t),t)$ is a three-dimensional, time-varying velocity field. Linear interpolation in space and time was used to obtain $V(X(t),t)$ from hourly FVCOM output. The contribution of advection to the particle position was updated by integrating equation (1) using an explicit fourth-order Runge-Kutta scheme with a time step of 600 seconds.

The LPM uses a random-walk process (Milstein scheme, Gräwe et al., 2011 and 2012; Rowe et al., 2016) to account for vertical turbulent mixing and *Microcystis* buoyant movement (Eq. 2). This is done after the calculation of particles' 3-D advective movement with ambient currents (Eq. 1).

$$z(\ +\delta\ ) = z(\ ) + W_b\delta\ + 0.5\frac{dK}{dZ}(z(\ ))[\Delta W^2 + \delta\ ] + \Delta W\sqrt{2K(z(\ ))} \tag{2}$$

Where $z(\ )$ is the vertical position of the particle at time $\ $, δ is the vertical random walk time step, W_b is buoyant velocity, K is vertical turbulent diffusivity, ΔW is a random variable drawn from a Gaussian distribution with zero mean and standard deviation $\sqrt{\delta}$. The detailed description and evaluation of the Milstein scheme are presented in Rowe et al. (2016). Note that an option exists in the LPM code to apply a spatially-uniform horizontal random-walk diffusion process. However, we applied the LPM as described by Rowe et al. (2016) for the CHAB forecast, in which horizontal diffusion was turned off.

An advantage of the Lagrangian approach is that each particle can have individual properties, which allows for taking into account varying buoyant velocities among *Microcystis* colonies. Therefore, particles (*Microcystis* colonies) were assigned with buoyant velocities randomly sampled from a buoyant velocity distribution estimated from a measured *Microcystis* colony size distribution and reported relationship between colony size and buoyant velocity described in Rowe et al. (2016).

2.3 Eulerian tracer model (ETM)

The ETM resolves advective transport and turbulent mixing of tracer concentration (C) in the following advection-diffusion equation:

$$\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} + v\frac{\partial C}{\partial y} + w\frac{\partial C}{\partial z} - \frac{\partial}{\partial z}\left(k_h\frac{\partial C}{\partial z}\right) - \frac{\partial}{\partial x}\left(A_m\frac{\partial C}{\partial x}\right) - \frac{\partial}{\partial y}\left(A_m\frac{\partial C}{\partial y}\right) + w_b\frac{\partial C}{\partial z} = C_{source} - C_{sink} \quad (3)$$

where u, v, and w are the x, y, and z components of the water velocity, k_h is the vertical eddy diffusivity calculated using the Mellor and Yamada level 2.5 (MY-2.5) turbulent closure scheme (Mellor and Yamada, 1982), A_m is the horizontal diffusion coefficient calculated with the Smagorinsky eddy parameterization method (Smagorinsky, 1963), w_b is the buoyant velocity of *Microcystis* colonies, and C_{source} and C_{sink} represent the sources and sinks of *C* associated with biological processes. To represent the physical transport-only model and quantify the impact of physical processes on the bloom variability, the source and sink terms were turned off in this study.

In contrast to the LPM, where each particle can have an individual buoyant velocity randomly sampled from an estimated buoyant velocity distribution of *Microcystis* colonies, the Eulerian approach represents characteristics of the population mean instead of describing intrapopulation variability. Hence, a representative buoyant velocity of $9 \times 10^{-5} m/s$ was used in ETM, which represented 70% of measured buoyant velocities based on the frequency distribution histogram (Fig. S1). This provided the best model performance in the sensitivity analysis of buoyant velocities (supplementary material).

2.4 Property-carrying particle model (PCPM)

ETM is designed to solve the physical and biological processes in one equation (Eq. 3). This may become inefficient when running ensemble models with the need to test different configurations for biological components. This is because, in the ETM framework, any changes, regardless of physical or biological components, would require a time integration of the entire Eq. 3, even if the changes in biological terms do not impact physical processes (e.g., water movement and mixing). This creates an opportunity to save computing load by using an alternative approach to calculate advection and diffusion.

The particles in LPM represent chlorophyll mass, while the particles in Property-carrying particle model (PCPM) describe flow conditions and record associated environmental properties along the particle trajectories due to the water movement, namely "properties-carrying particles." These trajectories and recorded physical properties become the linking mechanism between the hydrodynamic and biological processes in the next step. It should be pointed out that the particles in LPM only appear in areas where algae are present because the particles represent chlorophyll mass. In contrast, the particles in PCPM cover the entire model domain to represent the water mass movement and associated environmental variables. In the region without algae, the recorded value of algae concentration is simply zero. Therefore, it is necessary to continue releasing particles according to the flow rates of the Detroit River in the PCPM to track the water entering the western basin of Lake Erie from the river. On the contrary, particles should not be released from the river in LPM as it is assumed the river carried no algae into the lake.

Second, the PCPM employs its own grid cell system with a rectilinear grid mesh (Fig. 2b) to calculate local average values of the particle-based environmental properties within each PCPM grid cell. Then, the cell-based state variables within a given PCPM grid cell are predicted in response to various biophysical processes (if they are included) by using the stored physical conditions and the initial conditions of the state variables (which is the information brought by particles from ambient locations in the previous step). This step is an Eulerian approach to simulating the biophysical process. After that, the cell-based state variables are reassigned to the particles within each PCPM grid cell, and these particles carry the updated properties of state variables to the next location. Note that no calculations are needed to calculate a particle's next-step destination, as particle trajectories have already been computed at the beginning. Such a cycle repeats at each time step.

PCPM differs from LPM as tracer particles in PCPM characterize 3-D fields of concentration encountered by the particles instead of representing chlorophyll mass associated with *Microcystis* colonies. As such, the PCPM separates the hydrodynamic transport and the remaining biophysical processes. Hence, the PCPM combines the pre-computed particle trajectories (Lagrangian approach) with local biophysical (in this case, vertical mixing and buoyancy) processes (Eulerian approach) to predict the spatial distribution and temporal variation of CHABs. In this PCPM implementation, 1.4 million initial particles were randomly distributed throughout the western basin of Lake Erie, within a total water volume of 38.6 km^3. The PCPM grid has a resolution of 1 km x 1km with 20 uniform sigma layers in vertical, each PCPM cell contains 15 particles on average. If no particles are present in a particular cell, PCPM uses the values from the previous time step. Additional particles are continuously introduced according to the flow rates of the Detroit River and the Maumee River to keep the same particle density as in the initial distribution (Xue et al., 2017 and 2018).

In this study, the PCPM first uses pre-recorded physical properties carried by particles to determine the PCPM grid cell-based average chlorophyll concentration ($\overline{Chl_{grid,n}}$) in a PCPM grid cell *n* as follows:

$$\overline{Chl_{grid,n}} = \sum_{j=1}^{L} Chl_{particle,m_j}/L \quad (4)$$

Where the summation includes all *L* particles (m_1, m_2,...m_L) currently within cell *n*. *L* is the total number of particles within that cell *n* and $Chl_{particle,m_j}$ is the concentration of chlorophyll associated with particle m_j. Then, the PCPM cell-based average chlorophyll concentration is updated through local processes (buoyant velocity and vertical mixing) in the local water column. The updated chlorophyll concentration is then redistributed from cells to particles (m_1, m_2,...m_L) to carry forward. The development, validation, and application of the PCPM are documented in Xue et al. (2017, 2018). Similar to the ETM, the PCPM does not resolve intrapopulation variability among *Microcystis* colonies. Thus, the PCPM used the same buoyant velocity as in the ETM.

2.5 Model initialization and Experiment design

To provide model initial conditions, we calculated cyanobacterial chlorophyll concentration based on the satellite-derived Cyanobacteria Index (*CI*), one of the European Space Agency's Sentinel-3 OLCI sensor products (https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.nodc:NOS-HABOFS-LakeErie, accessed Nov 8th, 2022). The CI has been used for detecting surface algal blooms and for quantitative mapping of cyanobacterial

chlorophyll concentrations in coastal oceans and inland waters (Binding et al., 2019). Chlorophyll was converted from the *CI* by an empirical relationship (Rowe et al., 2016; Tomlinson et al., 2016) as follows:

$$Chl = 12570 \times CI + 10 \quad (5)$$

We evaluated daily satellite images of *CI* for the three CHAB seasons in 2017-2019 and identified 134 images of CHAB occurrence to initialize the transport models, based on cloud cover conditions (Table 1; see below for details). Each 10-day model simulation was initialized from each image by assigning satellite-derived chlorophyll concentration to the transport model grids as the initial surface field using the nearest-neighbor interpolation. The missing data for those model grids under cloud cover were filled by model-predicted results on that day from the simulation initialized from the earlier image. Chlorophyll concentration can be directly used as the initial surface field for the ETM and PCPM. For the LPM, the chlorophyll concentration must be converted to an amount of Lagrangian particles by specifying a chlorophyll mass per particle and placing the corresponding number of particles in each model grid to represent the initial chlorophyll concentration. The initial vertical distribution of chlorophyll was determined by applying the surface chlorophyll concentration to the surface mixed layer (SML) depth, as described in Rowe et al. (2016).

The satellite images used in the model simulations were divided into four categories:

- No CHAB event
- CHAB event with clear sky
- CHAB event with low cloud coverage
- CHAB event with high cloud coverage

The "No CHAB event" category is defined as the case with no more than 5% of the no-cloud region having CHAB in the western basin of Lake Erie. The high and low cloud coverage categories are distinguished by whether more than 50% of the western basin is covered by clouds when there is a CHAB event. Finally, the clear sky category is those satellite images with no cloud cover that clearly show CHAB events in the western basin of Lake Erie. Model results of three transport models were evaluated against observations for the clear sky category and the low cloud coverage category.

Table 1. Description of satellite images for the 2017, 2018, and 2019 CHAB seasons

			Used for model evaluations		
	No CHAB event	CHAB event with high cloud coverage	CHAB event with low cloud coverage	CHAB event with clear sky	Total
2017	13	0	5	15	33
2018	26	15	1	2	44
2019	11	26	9	11	57
Total	21	36	27	50	134

To characterize the competition between vertical mixing, vertical advection, and buoyant velocity of *Microcystis* colonies, we used the dimensionless Péclet number (Pe), which is defined as the ratio of the advective transport rate due to the algal buoyant velocity and water vertical advection to the turbulent transport rate due to vertical eddy diffusivity:

$$Pe = \frac{(W_b + \overline{ww}) \times h}{\bar{k}} \quad (6)$$

Where W_b is the buoyant velocity (m/s), h is the water column depth (m), $\overline{ww}$ is the vertical advection velocity (m/s), and $\bar{k}$ is the column mean eddy diffusivity (m^2/s). The Péclet number in this study varied by orders of magnitude ($10^{-3} \sim 10^{3}$). $Pe < 1$ (10^{0}) indicates that vertical turbulent mixing dominates over buoyant velocity and vertical water advection, and vice versa.

2.6 Skill Assessments

Following the methods of Rowe et al. (2016) and Soontiens et al. (2019), we evaluated three transport models using the skill metrics of binary categorical variables and the mean absolute error (MAE) for a statistical assessment of model performance based on all 43 bloom events in three CHAB seasons. The binary categorical variables test whether models can correctly simulate the CHAB occurrence at each grid cell. Comparisons at each model grid cell were divided into four categories depending on whether or not a CHAB event is presented in a given grid. A CHAB event occurs on a grid cell if chlorophyll concentration is greater than 12 ug/L, which is a level-2 alert, defined by the World Health Organization, for short-term responses to toxic cyanobacteria in drinking water supplies, and a level-1 alert for monitoring and managing cyanobacteria in water bodies used for recreation (Chorus et al., 2021).

Four categories were defined: ***a*** (true positive): the model correctly predicts an observed CHAB event); ***b*** (false positive): the model predicts an unobserved event); ***c*** (false negative): the model fails to predict an observed event); ***d*** (true negative): both model and observation show no CHAB event). Two metrics, including Frequency Bias (FB) and Pierce Skill Score (PSS), were used to evaluate the model performance.

The FB is calculated as

$$FB = \frac{a+b}{a+c} \quad (7)$$

It gives the ratio of the number of grid cells over which the model predicts a CHAB event to the number of grid cells over which a CHAB event has actually been observed. FB >1 indicates an overestimation of the CHAB area (represented by the total number of grids where a CHAB event occurs) from the model, and vice versa.

The PSS is defined as

$$PSS = \frac{a}{a+c} - \frac{b}{b+d} \quad (8)$$

PSS compares the "true positive" and "false alarm" rates. PSS = 0 suggests the model predictions have equivalent rates of true positives and false alarms, which is the expectation of a random forecast. Therefore, a positive PSS score indicates that the model outperforms a random forecast, and a negative PSS score indicates that the model performs worse than a random forecast.

MAE measures the absolute difference between modeled chlorophyll concentration and observation.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|m_i - o_i| \quad (9)$$

Where n is the total number of grid cells. m_i and o_i are the modeled and observed chlorophyll concentration in each grid cell, respectively.

In addition, we evaluated the model performance against a persistence forecast, which assumes an observed CHAB pattern remains unchanged over time. The persistence forecast represents the best available information to forecast users in the absence of applicable models. By comparing the difference in PSS and MAE between transport models and against persistence forecasts, we characterized the quality of the additional information provided by the process-based transport models.

3. Results

3.1 Inter-day CHAB transport and variability

To examine the impacts of wind-driven transport, turbulent mixing, and buoyancy effect on CHAB surface variability, we presented detailed bloom evolution during two major CHAB events in this section (additional simulation results for low- and medium-size CHAB events are also provided in supplementary materials in Fig. S4-S9). A statistical assessment of model performance based on all 43 bloom events in three CHAB seasons is presented in Section 3.3. The most significant CHAB event occurred on September 22nd-23rd, 2017. On the 15th, an intense bloom originated from Maumee Bay and extended throughout the center of the western basin, as seen in the satellite image (Fig. 3a1). In the next eight days, the bloom was transported northeastward and reached the northern shore of western Lake Erie (Fig. 3a2, a3). During this period, the surface bloom significantly intensified. In the following days, the surface chlorophyll concentration decreased rapidly from September 26th through October 1st (Fig. 3a4, a5). Correspondingly, we conducted a continuous 15-day model simulation from September 15th to October 1st. The simulations were initialized from the satellite image on September 15th, and then the three models ran continuously until October 1st, so that the other satellite images represent independent observations that could be used for model assessment. All three transport models successfully captured the northeastward transport pattern of CHAB, as well as the bloom intensification during the first 8 days (Fig. 3a3-d3). All three transport models also successfully reproduced the following diminishing surface bloom with remnants near the northern shore on October 1st (Fig. 3b4-d4, b5-d5).

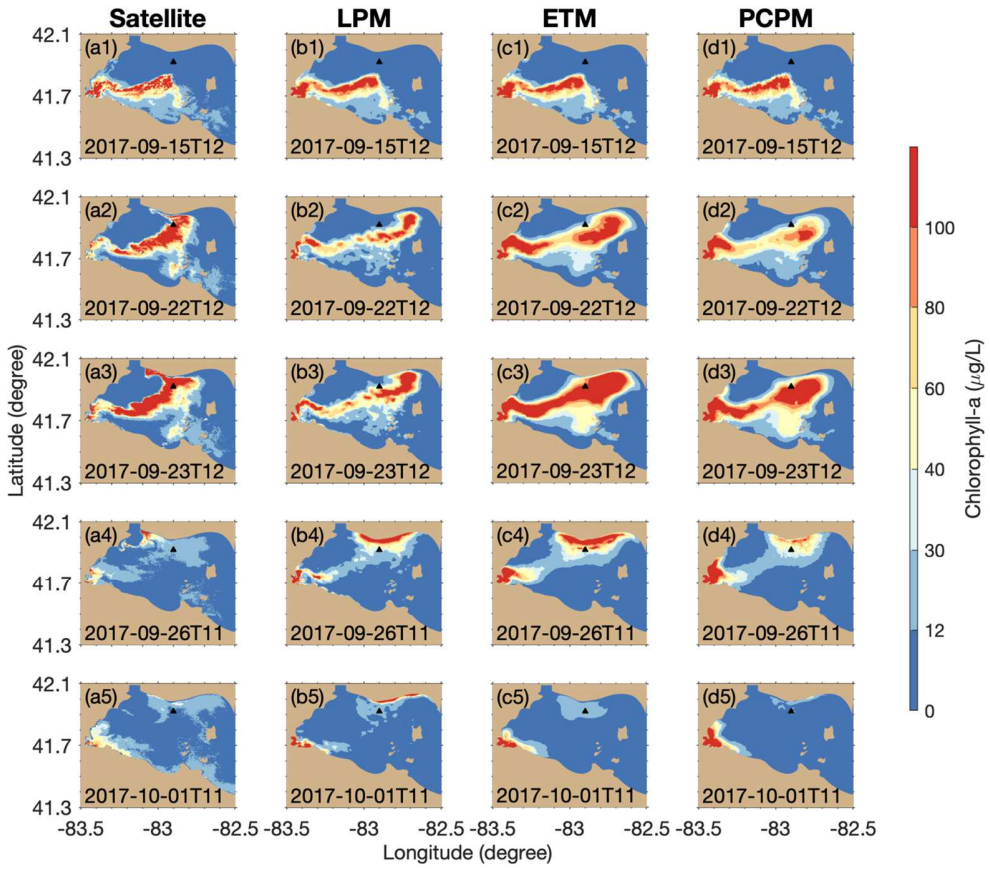

Fig. 3. Comparisons between simulations from three transport models (LPM, ETM, PCPM) and satellite-derived surface cyanobacterial chlorophyll concentration. Comparisons were made on the dates when satellite images were available during the event, including September 15th (a1, b1, c1, d1, initial fields), 22nd (a2, b2, c2, d2), 23rd (a3, b3, c3, d3), 26th (a4, b4, c4, d4), October 1st (a5, b5, c5, d5). The simulations were initialized from the satellite image a1 (September 15th) to create initial fields (b1, c1, d1), and then the models ran continuously until October 1st, so that the other satellite images represent independent observations that could be used for model assessment. The black triangle refers to the location for the time series plot of the Péclet number in Fig. 4(g) and the transect plot for chlorophyll concentration in Fig.5.

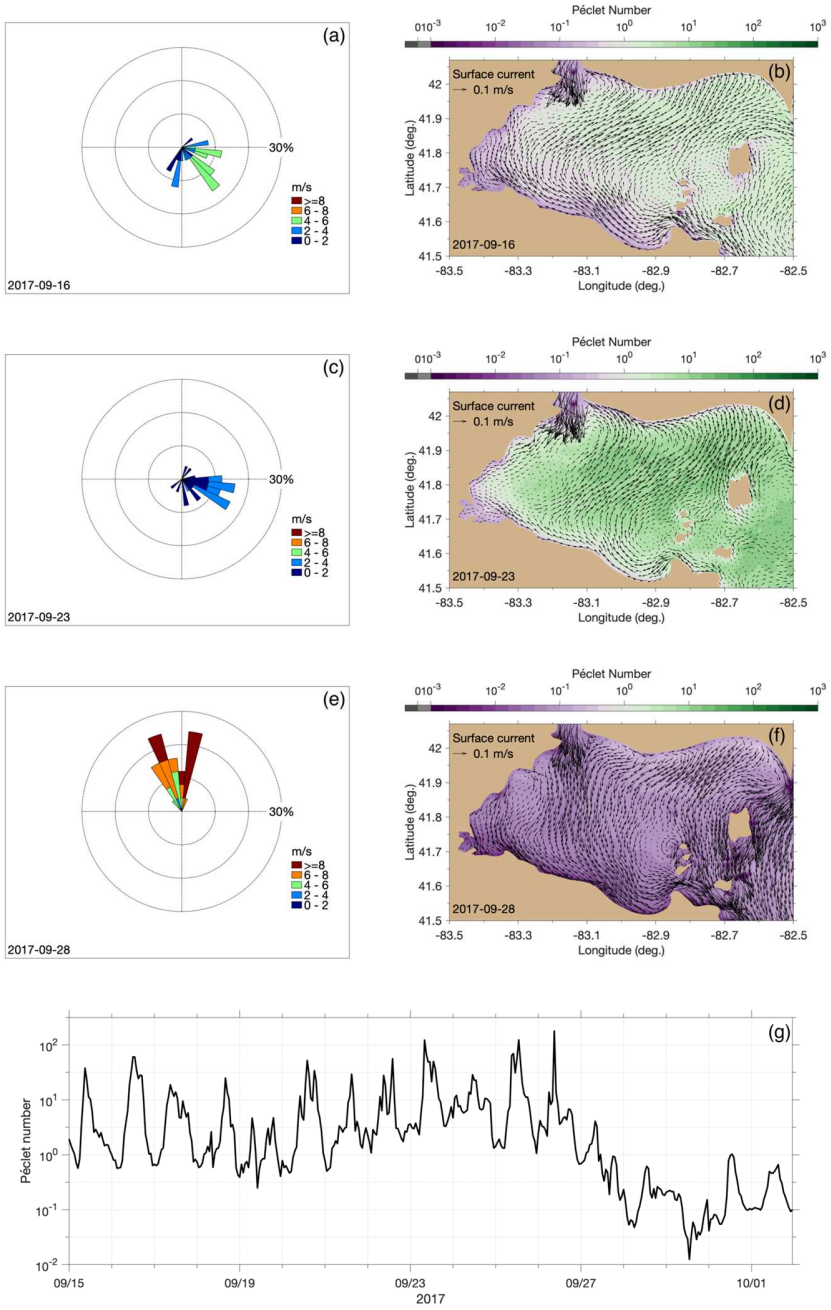

Fig. 4. Meteorological and hydrodynamic conditions during the CHAB event in 2017. Wind rose plots on September 16th, 23rd, and 28th, respectively (a, c, e). The wind rose plots were generated from the spatially averaged wind fields over the western basin that drove the hydrodynamic modeling. Modeled surface currents are overlaid with the Péclet number map on 16th, 23rd, and 28th, respectively (b, d, f). Time series plot of the Péclet number (g), at the location indicated as the black triangle in Fig. 3.

The fact that the models performed well in simulating not only how the bloom moved, but also its intensification and reduction, reinforces the predominant role of highly variable meteorological and hydrodynamic processes in short-term bloom evolution. Southeasterly wind prevailed during

September 16th–23rd (Fig. 4a). It generated northwestward currents in the shallow water along the south coast that primarily followed the wind direction (Fig. 4b). While in the deeper region of the western basin, the Ekman flow (surface currents turn right with respect to the wind forcing due to the Coriolis force) and the constraint of the shoreline boundary led to the northeastward transport (Fig. 4b). As a result, the bloom originated from Maumee Bay was transported to the northern shore of the western basin during this period. The intensification of the bloom on September 23rd was associated with the reduced vertical turbulent mixing, as reflected in the increase in the Péclet number. From September 16th to 23rd, the Péclet number in the center of western Lake Erie increased by more than an order of magnitude (Fig. 4b, d) due to the decreased wind speed from 4-6 m/s to 2-4 m/s (Fig. 4a, c). As a result, the algal buoyancy and water vertical advection gradually dominated over the vertical mixing, leading to upward vertical transport of algae from deeper water to the surface that intensified the surface bloom (Fig. 5a, b). The disappearance of blooms after September 23rd was primarily controlled by the enhanced vertical mixing due to the strong wind (Fig. 4e). The Péclet number decreased by three orders of magnitude throughout the western basin of Lake Erie during September 23rd-28th (Fig. 4d-f), reflecting the dominance of vertical mixing over buoyant velocity and water vertical advection, which mixed the algae throughout the water column and resulted in reduced surface chlorophyll concentration (Fig. 5c, d). The time evolution of bloom intensity was explained by the evolution of the Péclet number (Fig. 4g).

Furthermore, we separated the impact of algae buoyant velocity and water vertical advection on Péclet number at all grid points across the entire model domain. The results (Fig. S2) show that 95% of the vertical water velocities are below the buoyant velocity (9×10^{-5} m/s) used by the ETM and PCPM, and more importantly, 60% of the vertical water velocities are at least 10 times smaller than the buoyant velocity. Therefore, the Péclet number is primarily determined by the competition of the buoyant velocity and vertical turbulent mixing. As a result, we simplified the Péclet number (Pe_s) by excluding the impact of water vertical advection:

$$Pe_s = \frac{W_b \times h}{\bar{k}} \tag{10}$$

A comparison of Pe and Pe_s shows a very similar pattern (Fig. S3), which confirms that the buoyant velocity has the dominant influence over the vertical advection in Péclet number calculation. This is due to the fact that upward or downward water velocity, as resolved on a ~km-scale horizontal grid, has a much smaller magnitude compared to the buoyant velocity of *Microcystis* colonies (Fig. S1). In numerical fluid dynamic models, the sub-grid scale vertical velocities are represented by the turbulent diffusivity, which drives vertical mixing. Therefore, the Péclet number primarily represents buoyant velocity competing against turbulent mixing, and provides a useful indicator of conditions under which buoyant cyanobacterial colonies are likely to concentrate near the surface, versus being mixed through the water column.

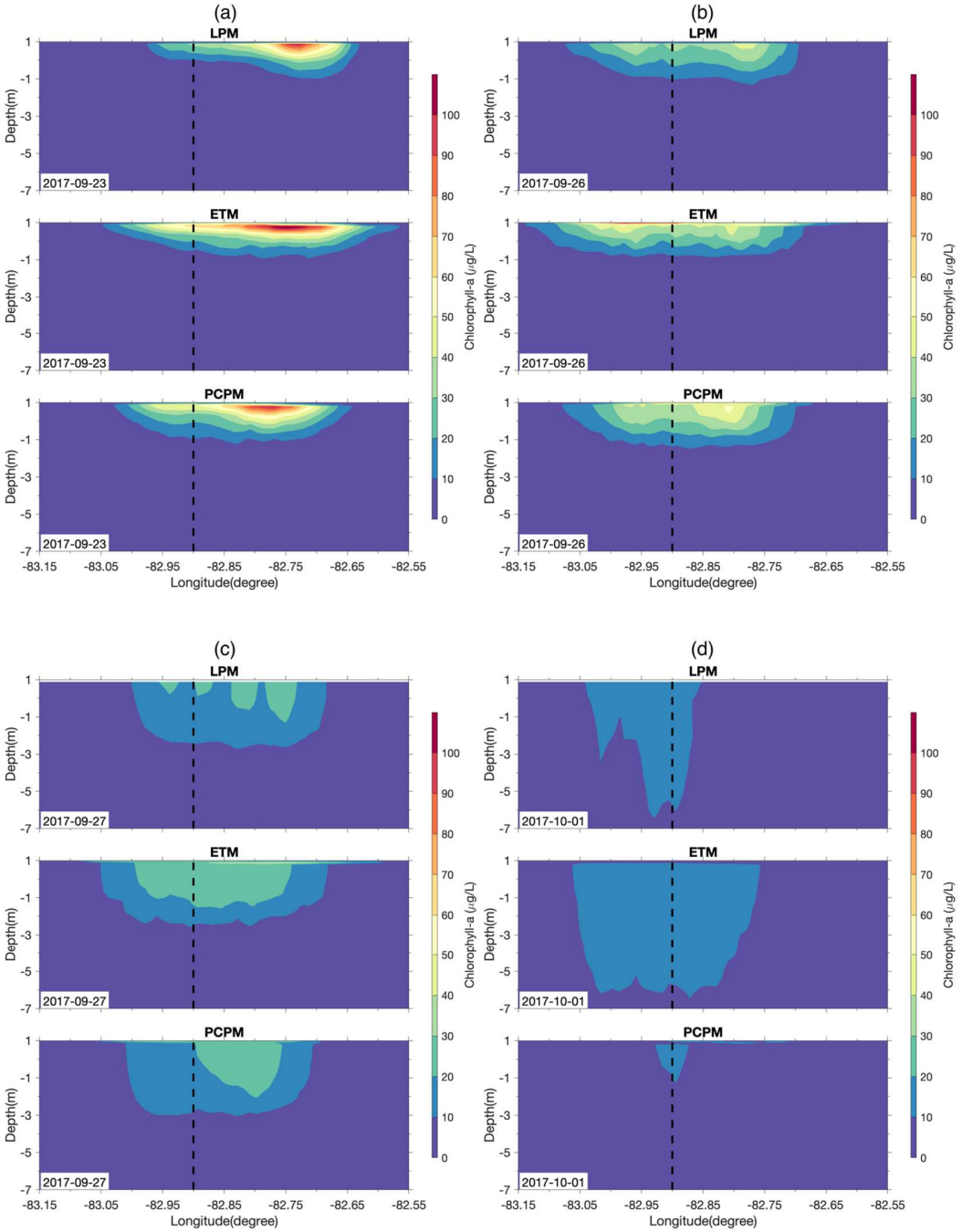

Fig. 5. Comparison of daily average chlorophyll concentration simulated by LPM, ETM, and PCPM in the west-east transect on a) September 23rd, b) September 26th, c) September 27th, d) October 1st across the black triangle location shown in Fig.3.

Similarly, in 2019, a severe CHAB event occurred from July 29th to August 5th (Fig. 6). On July 29th, the bloom with a surface chlorophyll concentration of 40-60 $\mu g/L$ occurred near the west shore (Fig. 6a1). Surface chlorophyll concentration increased to more than 100 $\mu g/L$ on July 30th.

The bloom moved eastward (Fig. 6a2) with two branches, one branch extending to the center of the western basin and the other one moving along the south coast. On August 2nd, the central branch of the bloom moved further eastward, leaving a "CHAB finger" pointing to the northern shore (Fig. 6a3). The finger extended to the north shoreline of the western basin, forming a semi-circular-shaped front when it met with the water mass from the Detroit River outflow from August 2nd to 5th (Fig. 6a4).

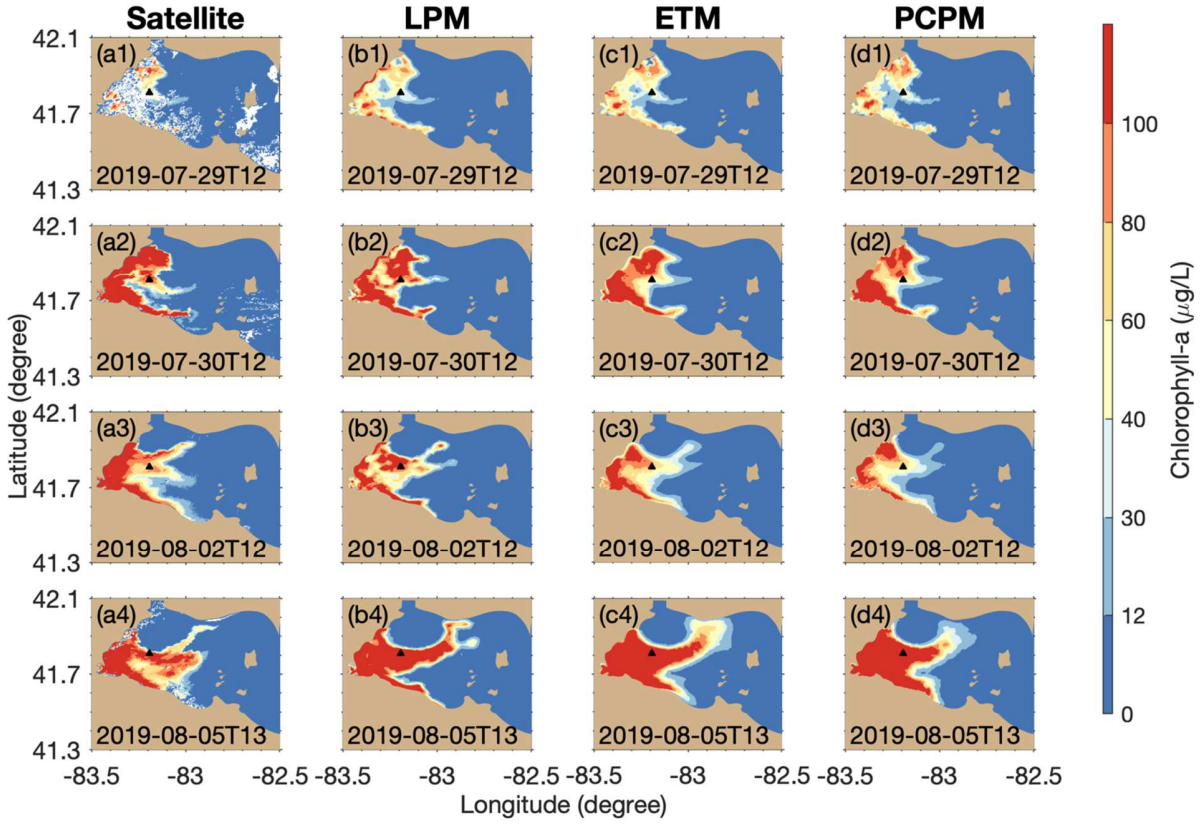

Fig. 6. Comparisons between simulations from three transport models (LPM, ETM, PCPM) and satellite-derived surface cyanobacterial chlorophyll concentration. Comparisons were made on the dates when satellite images were available during the event, including July 29th (a1, b1, c1, d1, initial fields), 30th (a2, b2, c2, d2), August 2nd (a3, b3, c3, d3), 5th (a4, b4, c4, d4). The simulations were initialized from the satellite image a1 (July 29th) to create initial fields (b1, c1, d1), and then the models ran continuously until August 5th, so that the other satellite images represent independent observations that could be used for model assessment. The black triangle refers to the location for the time series plot of the Péclet number in Fig. 7(e) and the transect plot for chlorophyll concentration in Fig.8.

All three transport models captured the northeastward propagation of the bloom and the semi-circular-shaped front around the Detroit River mouth. The three transport models also successfully reproduced the increase in surface chlorophyll concentration on July 30th, followed by a decrease of surface chlorophyll concentration on August 2nd and a re-intensification of chlorophyll

concentration on August 5th. This, again, highlights the significant impact of buoyancy and vertical mixing on surface CHAB intensity.

The formation of the "CHAB finger" and the semi-circular-shaped front resulted from wind-driven surface currents and the outflow from the Detroit River. A southwesterly wind of 6-8 m/s on July 29th (at the beginning of the event) (Fig. 7a) and southeasterly wind of 2-4 m/s (Fig. 7c) on August 5th (at the end of the event) favored the northeastward flow carrying high concentration of chlorophyll (Fig. 7b, d). Meanwhile, the water mass from the Detroit River flowed southward in the northern part of the basin with a counter-clockwise turn to the northeast to exit the western basin. Correspondingly, the chlorophyll front was formed between the two water masses.

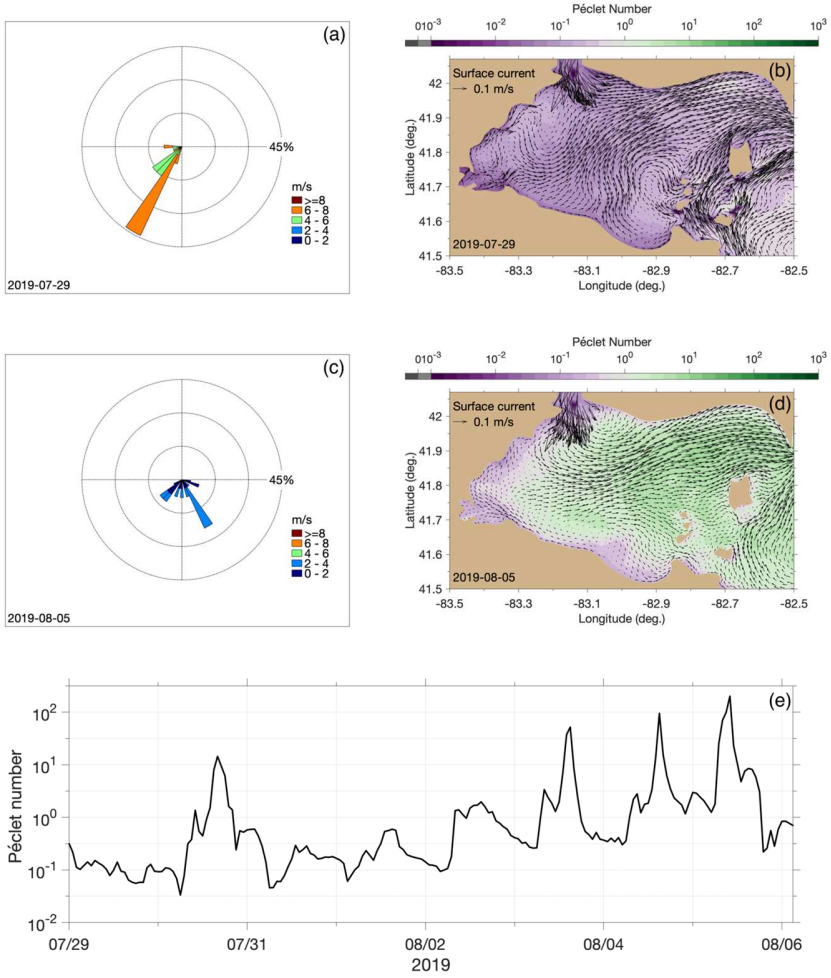

Fig. 7. Meteorological and hydrodynamic conditions during the CHAB event in 2019. Wind rose plots on July 29th and August 5th, respectively (a, c). The wind rose plots were generated from the spatially averaged wind fields over the western basin that drove the hydrodynamic modeling. Modeled surface currents are overlaid with the Péclet number map on July 29th and August 5th, respectively (b, d). Time series plot of the Péclet number (e), at the location indicated as the black triangle in Fig. 6.

The modeled vertical distribution of chlorophyll varied along with the changing Péclet number (Fig. 7 and 8). Offshore, where surface chlorophyll concentration varied significantly, the chlorophyll concentration of 50 $\mu g/L$ was uniformly distributed from the surface to the SML depth of 5 meters (Fig. 8a) on July 29th. The rapidly elevated Péclet number on July 30th resulted in the

first intensification of surface bloom (Fig. 7e, 6a2-d2). On August 2nd, buoyancy-driven surface bloom intensification extended to the adjacent area (83.1 °W to 83.3 °W) with horizontal transport and diffusion. A continued increase in Péclet number due to reduced wind and mixing since August 2nd led to a further decrease in SML depth and intensification of the surface bloom, with the chlorophyll concentration exceeding 100 $\mu g/L$ (Fig. 8c, d). The three transport models simulated the vertical structure of chlorophyll similarly, with some noticeable differences between the LPM and the other two models. This is mainly because the LPM used a distribution of buoyant velocity while the ETM and PCPM applied a uniform buoyant velocity, as discussed in Section 2.2-2.4.

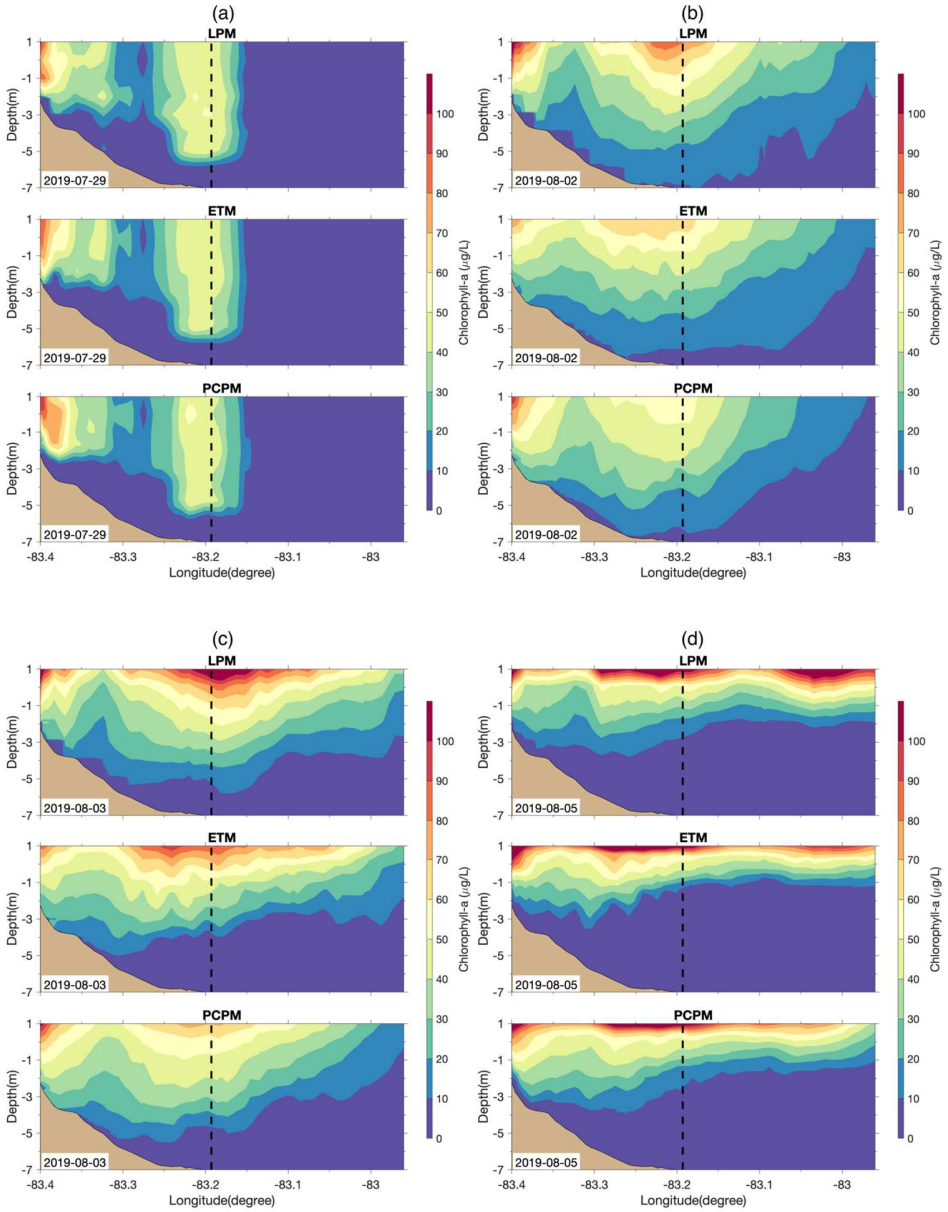

Fig. 8. Comparison of daily average chlorophyll concentration simulated by LPM, ETM, and PCPM in the west-east transect on a) July 29th, b) August 2nd, c) August 3rd, d) August 6th across the black triangle location shown in Fig.6.

3.2 Diel variability of CHAB concentration

Model results also revealed significant intraday variation of surface chlorophyll concentration in a diel cycle. Figure 9 presents the variation on August 5th, 2019 as a typical example. At 4:00 AM, surface chlorophyll concentration was around 40-60 $\mu g/L$ in the region (Figure 9a1, b1, c1). At 1:00 PM, the surface chlorophyll concentration increased significantly to 100 $\mu g/L$ with the CHAB area nearly unchanged (Figure 9a2, b2, c2). At 8:00 PM, surface chlorophyll concentration decreased again to 30-40 $\mu g/L$ (Figure 9a3, b3, c3). Such a cycle was clearly shown in the deeper places (water depth > 4 m) where water is more influenced by convective cooling. Surface chlorophyll concentration was low during the nighttime due to surface cooling-induced mixing and increased significantly when water was re-stratified (mixing weakened) during the daytime. Correspondingly, Fig. 10 shows that the Péclet number was greater than one from 4:00 AM to 12:00 PM, leading to algae floating upward from the deep layer to the surface and reaching its peak surface concentration at noon. The Péclet number became less than one afterward, consistent with the time when surface chlorophyll concentration started to decrease. All three transport models simulated the diel cycle, although PCPM and ETM showed a higher peak value and reached peak value slightly earlier than the LPM due to their different buoyancy configurations.

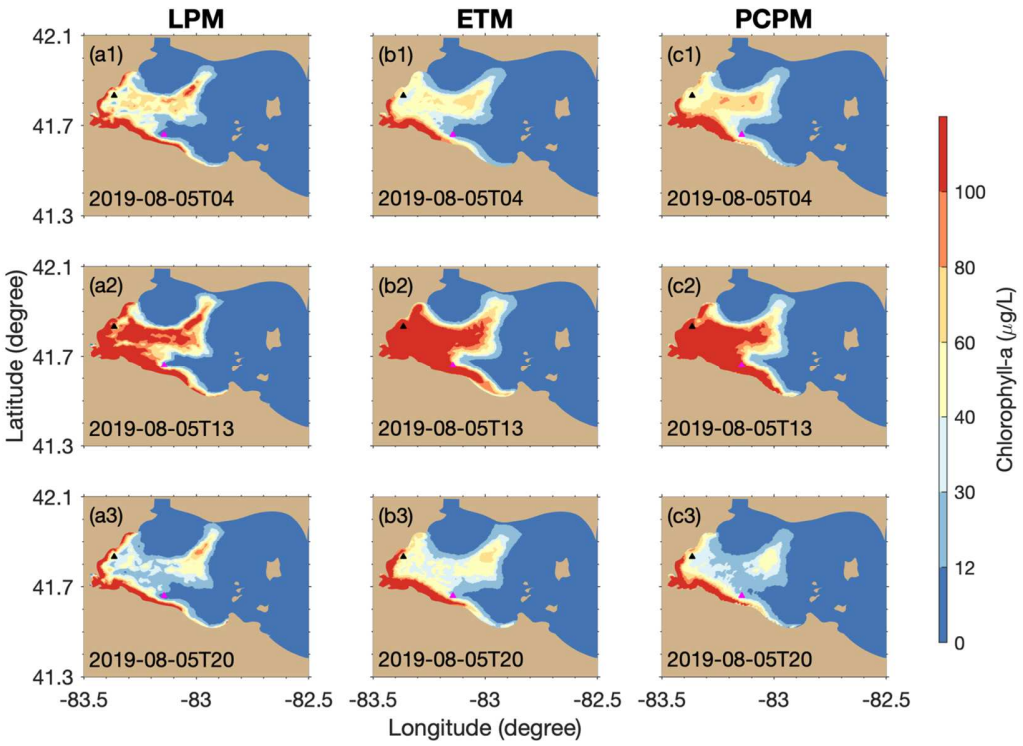

Fig. 9. Intraday variability of cyanobacterial chlorophyll concentration in response to diel heating and cooling of the surface, simulated by three transport models on August 5th, 2019. The black triangle represents the area of a high chlorophyll concentration, where the variation of surface chlorophyll concentration and Péclet number is shown in Fig. 10.

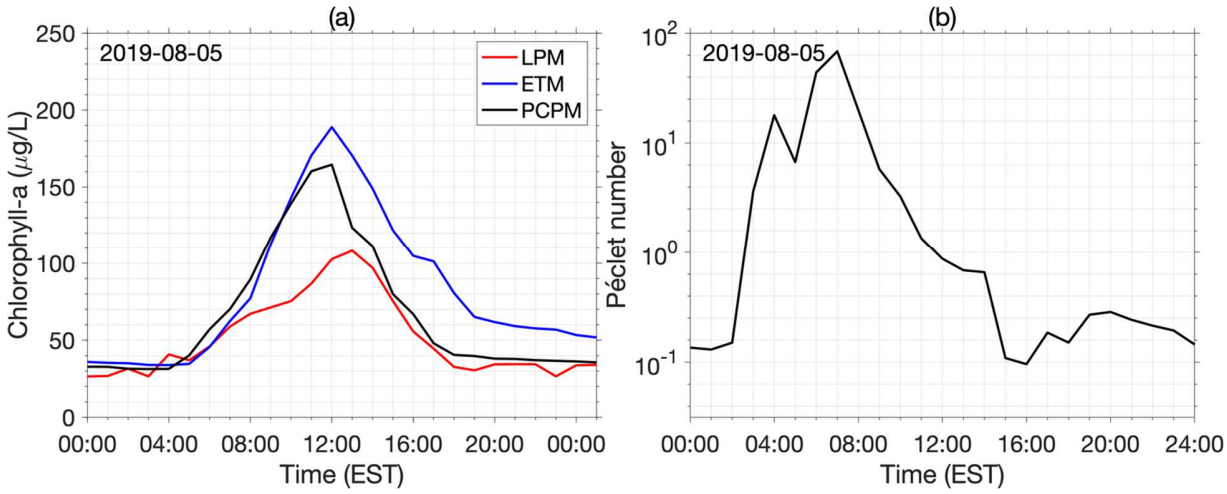

Fig. 10. Variation of simulated surface chlorophyll concentration (a) and Péclet number (b) at the black triangle marked in Fig. 9.

3.3 Model skill statistics

The two bloom events we presented in detail in previous sections are aimed at identifying the impact of physical processes on CHAB evolution. In this section, we focus on a statistical assessment of model performance based on all 43 bloom events in the three CHAB seasons (2017–2019) to evaluate the models' overall performance. We evaluated model performance from 24- to 240-hour hindcasts for each CHAB event in the three consecutive CHAB seasons and established model skill statistics to identify each model's strengths and limitations. The model skill assessments were grouped into two-day intervals based on model forecast length (Table 2). There were 103 model-satellite matchups for each simulation, with more than 10 for each group.

Table 2. Summary of available satellite images used for skill assessment, grouped into two-day intervals.

Forecast days	# satellite images available for model evaluation
1-2	15
3-4	31
5-6	12
7-8	32
9-10	13
	Total: 103

All three transport models had positive PSS scores, indicating that they have greater skill in capturing the occurrence of CHAB/non-CHAB events than a random forecast (Fig. 11). All three models' PSS scores decreased from (~0.75-0.6) as the number of simulation days increased, which

shows that the models' prediction accuracy decreased over longer prediction periods, as expected. The FB values for all three transport models were less than 1.0 (~0.7-0.9), indicating that all the models underestimated the CHAB area. Among these three models, the ETM had the best performance based on PSS-FB metrics.

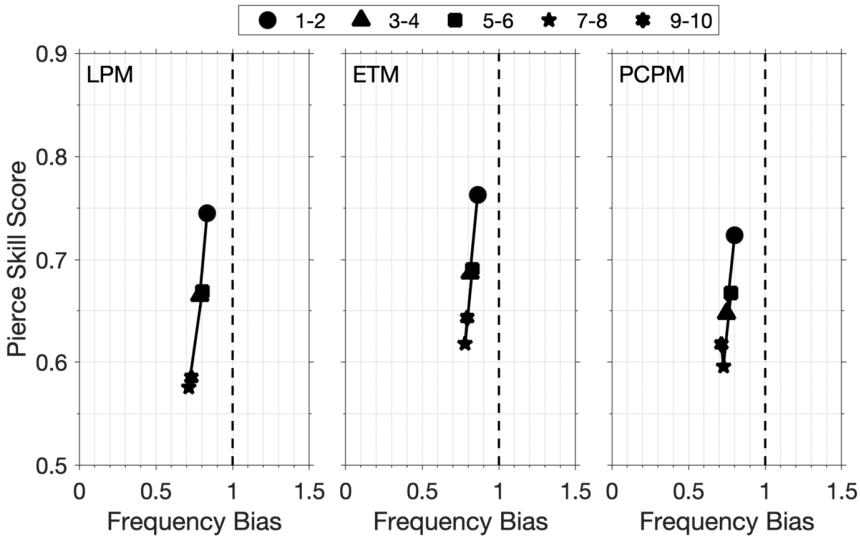

Fig. 11. PSS and FB plots for the LPM, ETM, and PCPM hindcast simulations. Markers indicated the number of forecast days since the model's initialization.

In addition, Fig. 12 evaluates the transport models' performance against a persistence forecast. As persistence forecast assumes a steady CHAB pattern over time, it performed well in 1-2 day prediction, capturing both CHAB/non-CHAB events and chlorophyll concentrations. This is because the CHABs in Lake Erie have several persistent features, which contribute to the skill of the persistence forecast. For example, CHABs often persist in the southern and western nearshore zones due to relatively weak currents and long residence time. Also, the bloom is rarely present in the Detroit River plume or beyond 82.7°W, as indicated by 13 years of Lake Erie CHAB spatial patterns compiled by Wynne and Stumpf (2015). The three transport models outperformed the persistence forecast in the following days, demonstrating the cumulative impacts of transport and mixing on bloom spatiotemporal distribution.

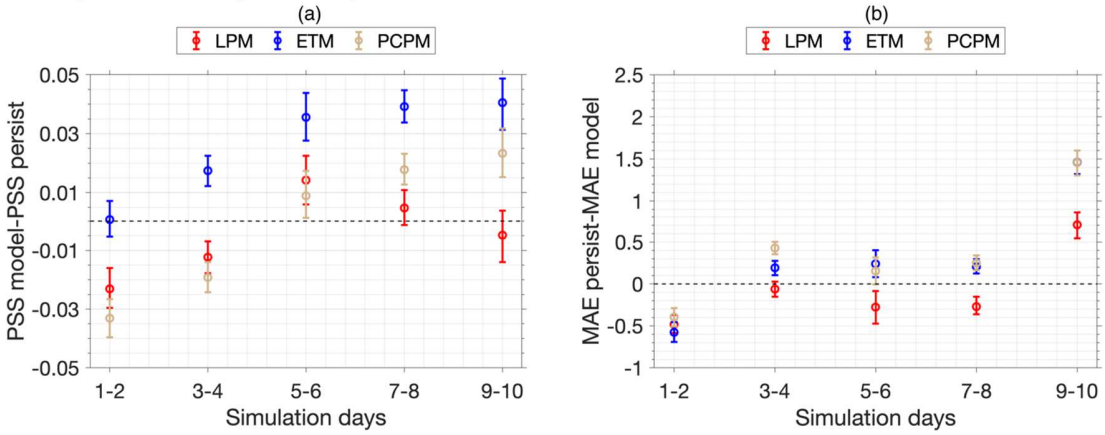

Fig. 12. The difference in skill scores between the three transport models (LPM, ETM, and PCPM) and the persistence forecast: a) PSS score difference. b) MAE score difference. Positive values

indicate a transport model performs better than the persistence forecast. Error bar represents the 95% confidence interval calculated using bootstrapping techniques (Hogan and Mason, 2012).

Among the three transport models, the ETM performs the best in terms of the PSS and FB metrics. With respect to the MAE metric, the ETM and PCPM show similar performances, which are slightly better than the performance of the LPM. Considering the differences in PSS and MAE scores among the three transport models are insignificant, it suggests that all three transport models have a similar level of skill and the ETM performs the best in the overall evaluation. Note that we also tested model sensitivity to buoyant velocities of *Microcystis* colonies, including the second experiment with a high buoyancy velocity of $18 \times 10^{-5}\ m/s$ (representing *Microcystis* colonies with large diameters) and the third experiment excluding buoyant velocities (supplementary materials). The buoyant velocity of $9 \times 10^{-5}\ m/s$ used here in the ETM and PCPM represented 70% of measured buoyant velocities based on the frequency distribution histogram of the measured buoyant velocities (Fig. S1) and provided the best model performance in our sensitivity analysis (Fig. S4).

4. Discussion and Conclusion

4.1 Importance of physical processes on CHAB forecast

For forecasting the short-term evolution of CHAB events, this study reveals the importance of highly dynamic flow patterns and associated transport and mixing in response to weather-scale wind events. While in the context of long-term mean circulation, Beletsky et al. (2013) summarized that the western basin circulation is driven by the Detroit River inflow moving eastward out of the basin, with wind modulating the circulation to a certain extent. However, short-term responses of flow patterns to weather-scale winds have a significant impact on CHAB forecasts. As CHABs are persistently located in Maumee Bay (Wynne and Stumpf 2015), wind-driven currents play the most critical role in the bloom movement to locations that are less commonly affected, which is the ultimate goal of what the transport models are intended to forecast. Our results show that, depending on wind direction, wind events can cause nearshore currents that transport CHABs eastward from Maumee Bay toward the Toledo water intake, northward toward the Monroe water intake, and long-distance northeastward transport to the Canadian shore (Fig. 3 and 6). Focusing on how the flow patterns quickly respond to wind events, our results provide additional insights in comparison to the previous understanding of the Detroit River as the main driver of currents in the western basin.

This study also demonstrated that the competition between vertical mixing and buoyant velocity (with a much less significant impact from vertical advection) is a key factor in determining the surface intensity and vertical distribution of CHABs. When the colony's buoyancy is strong enough to keep *Microcystis* concentrated within the SML, the chlorophyll concentration changes with SML depth (Rowe et al., 2016). Therefore, strong turbulent mixing keeps the algae distributed homogeneously within deeper SML, while surface algae can quickly intensify when vertical mixing is reduced (Fig. 5 and 8). Such a competition between vertical mixing and buoyant velocity is well expressed by the (simplified) Péclet number defined in this study (Fig. 10 and 11). The Péclet number serves as a relatively simple index that can predict the surface intensity and vertical distribution of CHAB. Previous studies also support these results. Wynne et al. (2010) show that the CHABs increased in both area and intensity for wind stress < 0.05 Pa and decreased for wind

599 stress > 0.1 Pa, and Fang et al. (2019) found that the average surface chlorophyll decreased by
600 about 6.2% with a wind speed increase by 1 m/s. In our study, buoyant velocity was held constant
601 as a simplifying assumption, so variation in Péclet number was largely driven by varying turbulent
602 diffusivity. However, other studies have shown that *Microcystis* colony buoyant velocity can vary
603 across lake systems (Den Uyl et al., 2021) and due to diel variation in cell carbohydrate content
604 (Medrano et al., 2013). Thus, by accounting for the main drivers of *Microcystis* vertical
605 distribution, buoyancy and turbulent diffusivity, Péclet number regarded as a more general
606 indicator of *Microcystis* vertical distribution across systems than, for example the effects of wind
607 speed, which would depend on local variables such as the exposure of a lake to wind and the
608 buoyancy of the local *Microcystis* population.

609 **4.2 Strengths and limitations of the transport models**

610 As predicting CHAB intensity and spatial distribution is critical to Lake Erie ecosystem
611 management, this study presents a comprehensive evaluation of Eulerian, Lagrangian, and hybrid
612 transport modeling approaches for Lake Erie CHAB hindcasts. To ensure an objective assessment,
613 these models were run in the most recent 3-D version and evaluated under identical biophysical
614 conditions and with the same statistical skill measures. The skill assessments show that three
615 transport models could achieve similar levels of prediction accuracy based on transport-only
616 simulations in this study.

617 However, in the context of general biophysical modeling, each of these three models has its own
618 advantages in different situations. While a known benefit of Lagrangian models is that they may
619 be less susceptible to numerical diffusion than Eulerian models, recent studies show that when
620 modeling CHAB with a high level of spatial detail, the spatial mismatch between modeled and
621 observed fields can cause a significant penalty in skill, and smoothing the spatial fields can
622 improve skill (Gill et al. 2017). Thus, smoother CHAB distributions in the Eulerian models (e.g.,
623 ETM and PCPM, which resolve spatiotemporally varying horizontal eddy diffusivity) may have
624 contributed to greater skill. Incorporating horizontal diffusion using the random walk method in
625 LPM has the potential to improve the model skill; however, we elected to implement the LPM in
626 this study as it had previously been applied in the Lake Erie HAB forecast. Even so, spatially-
627 varying horizontal diffusivity is more difficult to apply in an LPM due to potential numerical
628 artifacts (Ross et al., 2004), and ETM can represent more realistic spatially-varying horizontal
629 diffusivity (Smagorinsky et al., 1963).

630 ETM can better represent continuous fields of concentration than LPM. Nearly all lower-food web
631 biological models used in the Great Lakes represent nutrients, phytoplankton, zooplankton, and
632 detritus compartments as continuous fields and describe the biological process in the water column
633 in the Eulerian framework. Therefore, the lower-food web biological models can be directly
634 coupled to ETM due to their compatibility (Xue et al. 2014; Rowe et al. 2017). On the other hand,
635 LPM can represent properties that vary over a population, such as buoyant velocity, and track
636 exposure to environmental conditions over time, which is essential for individual-based models of
637 organisms (Li et al., 2014). PCPM, to some extent, possesses the advantages of ETM but with
638 greater computational efficiency. PCPM is 15 times faster than ETM using the same CPUs. PCPM
639 is also 30% faster than LPM. This can be a critical factor depending on how much computing
640 power is available and how many simulation scenarios an application requires.

Finally, although the present model configurations focus on simulating the impacts of physical processes on CHAB forecast, it is critical to include biological processes to resolve the source and sink processes of algae biomass in long-term biophysical simulations. The fact that ETM and PCPM performed as well as or better than the LPM sets up an alternative path to developing more biological realism in future models using Eulerian or hybrid approaches.

Acknowledgments

This is the contribution 96 of the Great Lakes Research Center at Michigan Technological University, and NOAA GLERL contribution number XX. This research was funded by the National Oceanic and Atmospheric Administration's National Centers for Coastal Ocean Science under award NA17NOS4780186. Liu was funded by the award to the University of North Carolina at Wilmington through NOAA (1305M320PNRMA0357). The Michigan Tech high performance computing cluster, *Superior*, was used in obtaining the modeling results presented in this publication.

Reference

1. Beletsky, D., Schwab, D. J., Roebber, P. J., McCormick, M. J., Miller, G. S., & Saylor, J. H. (2003). Modeling wind-driven circulation during the March 1998 sediment resuspension event in Lake Michigan. *Journal of Geophysical Research: Oceans*, *108*(C2).
2. Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., ... & Manikin, G. S. (2016). A North American hourly assimilation and model forecast cycle: The Rapid Refresh. Monthly Weather Review, 144(4), 1669-1694.
3. Baker, D. B., Confesor, R., Ewing, D. E., Johnson, L. T., Kramer, J. W., & Merryfield, B. J. (2014). Phosphorus loading to Lake Erie from the Maumee, Sandusky and Cuyahoga rivers: The importance of bioavailability. Journal of Great Lakes Research, 40(3), 502-517.
4. Beletsky, D., Hawley, N., & Rao, Y. R. (2013). Modeling summer circulation and thermal structure of Lake Erie. *Journal of Geophysical Research: Oceans*, *118*(11), 6238-6252.
5. Binding, C. E., Zastepa, A., & Zeng, C. (2019). The impact of phytoplankton community composition on optical properties and satellite observations of the 2017 western Lake Erie algal bloom. *Journal of Great Lakes Research*, *45*(3), 573-586.
6. Chorus, I., & Welker, M. (2021). Toxic cyanobacteria in water: a guide to their public health consequences, monitoring and management (p. 858). Taylor & Francis.
7. Huret, M., J. Runge, C. Chen, G. Cowles, Q. Xu, and J. Pringle (2007), Dispersal modeling of fish early life stages: Sensitivity with application to Atlantic cod in the western Gulf of Maine, *Mar. Ecol. Prog. Ser.*, 347, 261–274
8. Fang, S., Del Giudice, D., Scavia, D., Binding, C. E., Bridgeman, T. B., Chaffin, J. D., ... & Obenour, D. R. (2019). A space-time geostatistical model for probabilistic estimation

of harmful algal bloom biomass and areal extent. *Science of the Total Environment*, *695*, 133776.

9. Gill, D., Ming, T., & Ouyang, W. (2017). Improving the Lake Erie HAB Tracker: A Forecasting & Decision Support Tool for Harmful Algal Blooms (Doctoral dissertation).
10. Gilleland, E. (2010). Confidence Intervals for Forecast Verification: Practical Considerations. *National Center for Atmospheric Research: Boulder, CO, USA*.
11. Henrichs, D. W., Hetland, R. D., & Campbell, L. (2015). Identifying bloom origins of the toxic dinoflagellate Karenia brevis in the western Gulf of Mexico using a spatially explicit individual-based model. Ecological modelling, 313, 251-258.
12. Hogan, R. J., and I. B. Mason (2012). Deterministic forecasts of binary events, in Forecast Verification: A Practitioner's Guide in Atmospheric Science, 2nd ed., pp. 31–59, John Wiley & Sons, Ltd., West Sussex, U. K.
13. Huret, M., J. Runge, C. Chen, G. Cowles, Q. Xu, and J. Pringle (2007), Dispersal modeling of fish early life stages: Sensitivity with application to Atlantic cod in the western Gulf of Maine, *Mar. Ecol. Prog. Ser.*, 347, 261–274
14. Li, Y., Chen, X., Chen, C., Ge, J., Ji, R., Tian, R., ... & Xu, L. (2014). Dispersal and survival of chub mackerel (Scomber Japonicus) larvae in the East China Sea. *Ecological Modelling*, *283*, 70-84.
15. Liu, Q., Rowe, M. D., Anderson, E. J., Stow, C. A., Stumpf, R. P., & Johengen, T. H. (2020). Probabilistic forecast of microcystin toxin using satellite remote sensing, *in situ* observations and numerical modeling. *Environmental Modelling & Software*, *128*, 104705.
16. Kane, D. D., Conroy, J. D., Richards, R. P., Baker, D. B., & Culver, D. A. (2014). Re-eutrophication of Lake Erie: Correlations between tributary nutrient loads and phytoplankton biomass. *Journal of Great Lakes Research*, *40*(3), 496-501.
17. Kelley, J. G. W., Chen, Y., Anderson, E. J., Lang, G. A., & Xu, J. (2018). Upgrade of NOS Lake Erie Operational Forecast System (LEOFS) to FVCOM: Model development and hindcast skill assessment.
18. Medrano, E. A., Uittenbogaard, R. E., Pires, L. D., Van De Wiel, B. J. H., & Clercx, H. J. H. (2013). Coupling hydrodynamics and buoyancy regulation in Microcystis aeruginosa for its vertical distribution in lakes. *Ecological Modelling,* 248, 41-56.
19. Mellor, G. L., & Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Reviews of Geophysics*, 20(4), 851-875
20. Nakamura, T., Adachi, Y., & Suzuki, M. (1993). Flotation and sedimentation of a single *Microcystis* floc collected from surface bloom. Water Research, 27(6), 979-983.
21. Rowe, M. D., Anderson, E. J., Wynne, T. T., Stumpf, R. P., Fanslow, D. L., Kijanka, K., ... & Davis, T. W. (2016). Vertical distribution of buoyant *Microcystis* blooms in a Lagrangian particle tracking model for short-term forecasts in Lake Erie. *Journal of Geophysical Research: Oceans*, *121*(7), 5296-5314.

720 22. Rowe, M. D., Anderson, E. J., Beletsky, D., Stow, C. A., Moegling, S. D., Chaffin, J.
721 D., ... & Ackerman, J. D. (2019). Coastal upwelling influences hypoxia spatial patterns
722 and nearshore dynamics in Lake Erie. *Journal of Geophysical Research: Oceans*, *124*(8),
723 6154-6175.

724 23. Ross, O. N., and J. Sharples (2004), Recipe for 1-D Lagrangian particle tracking models
725 in space-varying diffusivity, *Limnol. Oceanogr. Methods*, 2(9), 289–302.

726 24. Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I.
727 The basic experiment. *Monthly weather review*, 91(3), 99-164.

728 25. Steffen, M. M., Davis, T. W., McKay, R. M. L., Bullerjahn, G. S., Krausfeldt, L. E.,
729 Stough, J. M., ... & Wilhelm, S. W. (2017). Ecophysiological examination of the Lake
730 Erie *Microcystis* bloom in 2014: linkages between biology and the water supply
731 shutdown of Toledo, OH. Environmental science & technology, 51(12), 6745-6755.

732 26. Soontiens, N., Binding, C., Fortin, V., Mackay, M., & Rao, Y. R. (2019). Algal bloom
733 transport in Lake Erie using remote sensing and hydrodynamic modelling: sensitivity to
734 buoyancy velocity and initial vertical distribution. *Journal of Great Lakes Research*,
735 *45*(3), 556-572.

736 27. Stumpf, R. P., Johnson, L. T., Wynne, T. T., & Baker, D. B. (2016). Forecasting annual
737 cyanobacterial bloom biomass to inform management decisions in Lake Erie. *Journal of*
738 *Great Lakes Research*, *42*(6), 1174-1183.

739 28. Tomlinson, M. C., Stumpf, R. P., Wynne, T. T., Dupuy, D., Burks, R., Hendrickson, J., &
740 Fulton III, R. S. (2016). Relating chlorophyll from cyanobacteria-dominated inland
741 waters to a MERIS bloom index. Remote Sensing Letters, 7(2), 141-149.

742 29. Den Uyl, P. A., Harrison, S. B., Godwin, C. M., Rowe, M. D., Strickler, J. R., &
743 Vanderploeg, H. A. (2021). Comparative analysis of *Microcystis* buoyancy in western
744 Lake Erie and Saginaw Bay of Lake Huron. Harmful Algae, 108, 102102.

745 30. Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., & Dyble, J. (2010). Characterizing a
746 cyanobacterial bloom in western Lake Erie using satellite imagery and meteorological
747 data. *Limnology and Oceanography*, *55*(5), 2025-2036.

748 31. Wynne, T. T., Stumpf, R. P., Tomlinson, M. C., Fahnenstiel, G. L., Dyble, J., Schwab, D.
749 J., & Joshi, S. J. (2013). Evolution of a cyanobacterial bloom forecast system in western
750 Lake Erie: development and initial evaluation. *Journal of Great Lakes Research*, *39*, 90-
751 99.

752 32. Watson, S. B., Miller, C., Arhonditsis, G., Boyer, G. L., Carmichael, W., Charlton, M.
753 N., ... & Wilhelm, S. W. (2016). The re-eutrophication of Lake Erie: Harmful algal
754 blooms and hypoxia. Harmful algae, 56, 44-66.

755 33. Wynne, T. T., & Stumpf, R. P. (2015). Spatial and temporal patterns in the seasonal
756 distribution of toxic cyanobacteria in western Lake Erie from 2002–2014. *Toxins*, *7*(5),
757 1649-1663.

758 34. Xue, P., Eltahir, E.A.B., Malanotte-Rizzoli, P., and Wei, J. (2014), Local feedback
759 mechanisms of the shallow water region around the Maritime Continent, *J. Geophys. Res.*
760 *Oceans, 119, 6933–6951.*doi: 10.1002/2013JC009700

35. Xue, P., Schwab, D. J., Sawtell, R. W., Sayers, M. J., Shuchman, R. A., & Fahnenstiel, G. L. (2017). A particle-tracking technique for spatial and temporal interpolation of satellite images applied to Lake Superior chlorophyll measurements. *Journal of Great Lakes Research*, *43*(3), 1-13.

36. Xue, P., Schwab, D. J., Zhou, X., Huang, C., Kibler, R., & Ye, X. (2018). A hybrid Lagrangian-Eulerian particle model for ecosystem simulation. *Journal of Marine Science and Engineering*, *6*(4), 109.

37. Xue, P., Malanotte-Rizzoli, P., Wei, J., and Eltahir, E. A. (2020), Coupled Ocean-Atmosphere Modeling over the Maritime Continent: A Review. *Journal of Geophysical Research-Oceans*,125,e2019JC014978. ; DOI:10.1029/2019JC014978

38. Xue, P., Ye, X., Pal, J. S., Chu, P. Y., Kayastha, M. B., and Huang, C. (2022). Climate Projections over the Great Lakes Region: Using Two-way Coupling of a Regional Climate Model with a 3-D Lake Model. *Geosci. Model Dev.*, 15, 4425–4446. https://doi.org/10.5194/gmd-15-4425-2022.

39. Ye, X., Chu, P. Y., Anderson, E. J., Huang, C., Lang, G. A., & Xue, P. (2020). Improved thermal structure simulation and optimized sampling strategy for Lake Erie using a data assimilative model. *Journal of Great Lakes Research*, *46*(1), 144-158.