JGR Atmospheres

RESEARCH ARTICLE
10.1029/2024JD041794

Depth Matters: Lake Bathymetry Selection in Numerical Weather Prediction Systems

James Kessler[1], Eamon Espey[2], Alexander VanDeWeghe[3], Andrew D. Gronewold[3], Troy Sorensen[4], Bahram Khazaei[5,6], Eric James[7], Tatiana G. Smirnova[7,8], Matt Casali[9], David Yates[9], Nina Omani[9], John G. W. Kelley[5], Michael Barlage[10], Stanley G. Benjamin[7,8], and Eric J. Anderson[4]

[1]NOAA Great Lakes Environmental Research Lab, Ann Arbor, MI, USA, [2]North Carolina State University, Raleigh, NC, USA, [3]School for Environment and Sustainability, University of Michigan, Ann Arbor, MI, USA, [4]Colorado School of Mines, Golden, CO, USA, [5]NOAA Coast Survey Development Lab, Silver Spring, MD, USA, [6]Ocean Associates, Inc., Arlington, VA, USA, [7]NOAA Global Systems Lab, Boulder, CO, USA, [8]Cooperative Institute for Research in Environmental Sciences, Boulder, CO, USA, [9]National Center for Atmospheric Research, Boulder, CO, USA, [10]NOAA Environmental Modeling Center, College Park, MD, USA

Key Points:

- Lake bathymetry plays a critical role in determining lake surface water temperature in numerical weather models
- Changing bathymetry in the 1-D lake model of an operational weather model had strong effects on lake surface temperature for some lakes
- Lake-wide, daily average water surface temperature changed by as much as 10°C; on average root-mean-square deviation was 1°C

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. Kessler,
James.Kessler@NOAA.gov

Citation:

Kessler, J., Espey, E., VanDeWeghe, A., Gronewold, A. D., Sorensen, T., Khazaei, B., et al. (2025). Depth matters: Lake bathymetry selection in numerical Weather prediction systems. *Journal of Geophysical Research: Atmospheres*, *130*, e2024JD041794. https://doi.org/10.1029/2024JD041794

Received 17 JUN 2024
Accepted 20 NOV 2024

© 2025. The Author(s).
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract Lake surface conditions are critical for representing lake-atmosphere interactions in numerical weather prediction. The Community Land Model's 1-D lake component (CLM-lake) is part of NOAA's High-Resolution Rapid Refresh (HRRR) 3-km weather/earth-system model, which assumes that virtually all the two thousand lakes represented in CONUS have distinct (for each lake) but spatially uniform depth. To test the sensitivity of CLM-lake to bathymetry, we ran CLM-lake as a stand-alone model for all of 2019 with two bathymetry data sets for 23 selected lakes: the first had default (uniform within each lake) bathymetry while the second used a new, spatially varying bathymetry. We validated simulated lake surface temperature (LST) with both remote and in situ observations to evaluate the skill of both runs and also intercompared modeled ice cover and evaporation. Though model skill varied considerably from lake to lake, using the new bathymetry resulted in marginal improvement over the default. The more important finding is the influence bathymetry has on modeled LST (i.e., differences between model simulations) where lake-wide LST deviated as much as 10°C between simulations and individual grid cells experienced even greater departures. This demonstrates the sensitivity of surface conditions in atmospheric models to lake bathymetry. The new bathymetry also improved lake depths over the (often too deep) previous value assumed for unknown-depth lakes. These results have significant implications for numerical weather prediction, especially in regions near large lakes where lake surface conditions often influence the state of the atmosphere via thermal regulation and lake effect precipitation.

Plain Language Summary In order to make accurate and meaningful predictions, weather models must represent the land surface below; this includes the surface of lakes, which can have a significant impact on weather phenomena such as lake-effect precipitation or a shift in air temperature downwind of the lake. The state of the art weather model used by NOAA's National Weather Service represents lake depth in a somewhat unrealistic way–it assumes that the depth in a given lake is the same everywhere, and it often assumes lakes are too deep. This can lead to unrealistic surface conditions and potentially contribute to less accurate weather forecasts for North America. Replacing the existing lake depth with new, more realistic lake depth data, resulted in considerable changes in lake temperature and ice cover. Though this may not necessarily *improve* the atmospheric model performance, the large difference caused by a change in lake depth, motivates the need for more research on this topic. More specifically, are we using the right lake depth data to get the best possible weather forecasts?

1. Introduction

Lakes around the globe can have a profound impact on their surrounding meteorological conditions and climate, and those impacts vary depending on (among other factors) a lake's surface area, depth, orientation, and latitude (Notaro et al., 2015; Wright et al., 2013; K. Xiao et al., 2018). These impacts (also referred to as lake effects) reflect the high heat capacity of water and the generally higher (relative to over land) latent heat fluxes over water (Novick et al., 2016; Spence et al., 2013). Correctly representing these features and processes in numerical weather prediction (NWP) models, especially as spatial resolution continues to increase, is critical to continually

improving the integrity of these models and subsequently protecting human and ecological health, and property (Benjamin et al., 2022; Choulga et al., 2019). Some climate studies have used lake models, (e.g., Briley et al., 2021; Minallah & Steiner, 2021), but lakes have not been represented in operational NWP models at the National Oceanic and Atmospheric Administration (NOAA) with one recent exception: the NOAA operational High-Resolution-Rapid Refresh (HRRR) model (Benjamin et al., 2016; Dowell et al., 2022). HRRR uses a 1-D lake model from the Community Land Model (hereafter CLM-lake) v4.5 (Lawrence et al., 2019) with hourly cycling of lake surface temperature (LST) and ice cover (Benjamin et al., 2022) to improve heat and moisture fluxes near lakes. This complements the use of the RUC Land-Surface Model (Smirnova et al., 2016) for land points in the HRRR domain. The implementation of CLM-lake model in the HRRR model—hereafter HRRR-CLM-lake, and its initialization are described by Benjamin et al. (2022). The CLM-lake model is based on the Lake, Ice Snow and Sediment Simulator, or LISSS which was specifically developed to improve lake representation in atmospheric models (Subin et al., 2012). The configuration and performance of CLM-lake are both critical to how lakes are represented in the operational HRRR NWP model, and ultimately how key lake effects are propagated through numerical models into operational weather forecasts.

The utilization of a 1-D lake model without any horizontal transport or mixing in operational numerical weather prediction contradicts the common understanding that hydrodynamic and thermodynamic processes governing lake physics occur across both vertical and horizontal scales (Knauss & Garfield, 2016; Martynov et al., 2010). There has therefore been considerable research on understanding the extent to which parallel one-dimensional lake models represent these processes, and whether that representation propagates into varying levels of numerical weather prediction skill (Zhang et al., 2019). C. Xiao et al. (2016), for example, applied CLM-lake to the Laurentian Great Lakes and found it produced early seasonal cooling in the fall, followed by excessive lake ice cover and early seasonal warming in the spring. A similar study (Gu et al., 2013) found CLM-lake performed better for the shallowest of the Great Lakes (Lake Erie; mean depth ≈20 m) and (relatively) worse for the deepest (Lake Superior; mean depth ≈150 m). These results reinforce the logical notion that a lake's bathymetry, and its representation in HRRR-CLM-lake, impact actual and simulated heat content and other lake physical parameters.

The default configuration of HRRR-CLM-lake represents 1,864 lakes in CONUS (defined here as the HRRR domain covering lower 48 US and southern Canada) with a constant (distinct for each lake but spatially homogeneous within each lake) depth, taken from the Global Lake Data Base (GLDB version 2, Kourzeneva et al., 2012), where values were often based on observed maximum depth. It is commonly understood, however, that the spatial distribution of bathymetry in a lake influences the vertical distribution of heat and, in turn, LST (Hostetler & Bartlein, 1990). More specifically, thermal mass in a given water column is directly proportional to the total water volume (i.e., depth times surface area). Therefore, a relatively shallow region of a lake would be expected to release and absorb heat more quickly than a comparatively deeper region in HRRR-CLM-lake. The rate of heating or cooling in shallow areas of an actual lake is slowed in reality by horizontal mixing, but this (and any other horizontal) process is not represented in 1-D lake models. This implies that lake bathymetry configuration for 1-D models may have different requirements than more realistic 3-D lake models (which are generally best suited by the most realistic bathymetry).

To improve LST accuracy and lake-effect snow prediction, the Laurentian Great Lakes are now represented in the operational version of HRRR using an adaptation of the three-dimensional finite-volume community ocean model (FVCOM) with spatially heterogeneous 3 arcsec bathymetry derived from the NOAA National Geophysical Data Center (Anderson et al., 2018; Fujisaki-Manome et al., 2020). As noted above, every other lake in CONUS (i.e., aside from the Great Lakes) is represented in HRRR by CLM-lake. For most other CONUS lakes, comparable high-resolution lake depth data is not consistently available and, even if it were, configuring and maintaining 3-D hydrodynamic models (such as FVCOM) to fully leverage its benefits would be a major task.

There is, therefore, a clear need to identify pathways for improving representation of CONUS lake bathymetry in HRRR or any other operational NWP model, and to understand if those improvements would propagate into improved numerical weather predictions. Here, we address this need by assessing (through both model intercomparisons, and validation against in situ and remote sensing data) the impacts of different bathymetric data sets for CONUS lakes in HRRR-CLM-lake.

2. Methods

2.1. Rationale and Methodological Overview

To understand the relative impacts of different lake bathymetry configurations in the CLM-lake model, we set up and ran three parallel simulations. The first simulation used the default constant-depth bathymetry from GLDBv2 (Kourzeneva et al., 2012). The second simulation replaced the default bathymetry with the GLOBathy data set (Khazaei et al., 2022), which required resampling the original, high-resolution bathymetry data onto the 3-km resolution grids of the HRRR-CLM-lake. A third simulation was conducted using a depth-censored version of GLOBathy in which the depth of all grid cells was restricted to be no less than 2 m. This modification was intended to prevent extreme horizontal distortion in grid cells that were very shallow: grid cells with a resampled (from GLOBathy) depth of less than 2 m could be extremely "thin" and thermodynamically volatile (especially given that each grid in the HRRR configuration has a fixed surface area of 9 km^2). These three cases are hereafter referred to as "flat-bottom" or "default," "GLOBathy," and "GLOBathy2m," respectively.

We initially executed the first two simulations within the HRRR-CLM-lake retrospective framework on the NOAA research and development high-performance computing (HPC) system for a 6-week simulation window from September through mid-October 2019. This system executes the HRRR-CLM-lake model across all of CONUS with full (two-way) coupling between the lake/land surface and atmosphere. Preliminary results, however (not shown), indicated that the computational demand of this approach far exceeded expectations for our research study. More specifically, we found that the wall-clock time to model simulation time ratio was close to 1. We therefore modified our approach to run all three simulations in an offline (i.e., one-way coupled) mode, and to simulate lake physics on the largest 23 lakes, by surface area, in CONUS for which we were able to find readily available in situ LST data for model validation (see Section 2.4 below for details). The offline runs were at least three orders of magnitude faster and an entire year of simulation could be conducted in approximately 13 hr with a very modest number of CPUs. This has implications for future studies, since many more simulations (using different configurations or for longer time periods) could be easily conducted, if desired. Only results from flat-bottom and GLOBathy cases are presented here, as there is only marginal *lake-wide* difference between GLOBathy and GLOBathy2m (18 of 23 lakes had RMSD < 0.1°C and max RMSD was <1.0°C). However, there are significant local changes to affected grid cells so enforcing a minimum depth should be considered in future work.

2.2. Model Simulations

We conducted offline model simulations using CLM-lake v4.5 to be consistent with versions used in operational practice. The lake model was configured with a 3,600 s timestep on the University of Michigan's HPC and simulations were conducted for 01 January 2018 through 31 December 2019 for the flatbottom, GLOBathy, and GLOBathy2m bathymetric configurations. Lake temperature was initialized at the start of the simulation using the CLM-lake default for a "cold start": surface water is set to land surface temperature (provided by the operational HRRR model); water below 50 m is set to temperature of maximum density (4°C); water between 0 and 50 m is set by a linear slope. After initializing, we also discarded the first 10 months of output from each simulation to account for model spin-up (Garnaud et al., 2022, used 12 months as spin for a similar study involving a different 1-D lake model). We focused our skill assessment on the 2019 calendar year because of the relative abundance of in situ observations of LST over that period (see Sorensen et al., 2024, for details). For simulation intercomparison, we also included November and December of 2018 to identify differences during the transitional period into winter, when lakes undergo rapid sensible and latent heat loss. For seasonal intercomparison, the period December 2018 through November 2019 was used such that meteorological seasons consisted of consecutive months (DJF, MAM, JJA, SON) and there were no duplicated months.

We used output from the operational version of HRRR in years 2018 and 2019 (HRRRv3, Dowell et al. (2022)) as the surface boundary conditions for our offline CLM-lake simulations. This data was accessed via the NOAA-managed google cloud bucket (Gowan et al., 2022) and contains HRRR output for multiple atmospheric variables required to run CLM-lake including short- and long-wave radiation, surface pressure, precipitation rate, specific humidity and air temperature at 2 m above surface, zonal and meridional wind fields at 10 m. The HRRR data obtained was mostly complete during the period of simulations, but 20 and 39 (non-consecutive) hours of data were missing from calendar years 2018 and 2019, respectively. This missing data represents a very small fraction (~0.3%) of the total data. We linearly interpolated to infill these missing values using available data

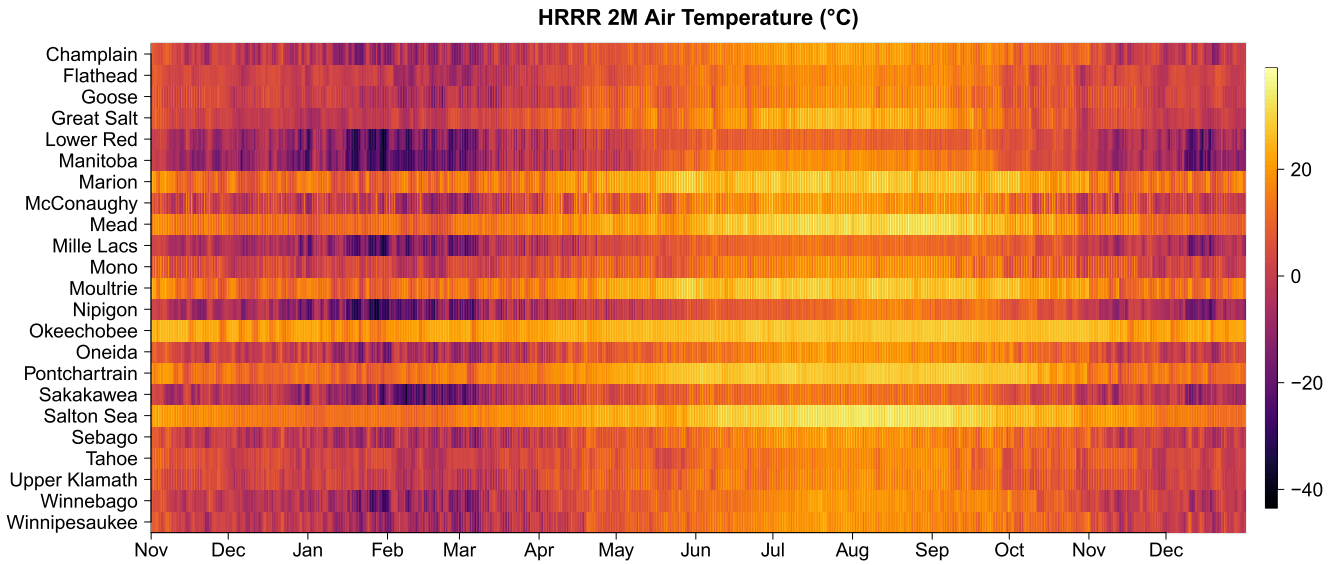

Figure 1. Hourly overlake (lake-wide average) 2 m air temperatures from the operational HRRR model from November 2018 through December 2019.

before and after the timestep of missing data. The CLM-lake model was forced by HRRR every hour and model output was also saved every hour. HRRR overlake 2 m air temperatures are shown for the 23 lakes studied in Figure 1.

2.3. Model Bathymetry Data

The default bathymetry encoded in the current operational version of the HRRR-CLM-lake model was developed specifically for use in numerical weather and climate models and originates from the GLDBv2 (Kourzeneva et al., 2012). We made no modifications to this data set in our CLM default (i.e., "flatbottom") simulations, but it is important to note that four lakes (Champlain, Winnipeg, St. Clair, Saint-Jean) were originally encoded in CLM-lake with a spatially variable bathymetry. Our results for Lake Champlain, therefore, will provide insight into whether the GLOBathy represents an improvement over the default spatially heterogeneous (and, perhaps, more accurate) bathymetry. Benjamin et al. (2022) (see their Section 3.3) pointed out that the assumed lake-depth value of 50 m for unknown lakes was far too large.

We then incorporated the GLOBathy (Khazaei et al., 2022) data set into our second and third model simulations. GLOBathy is a readily available model-based high-resolution (1 arcsec) data set that includes over 1.4 million lakes globally, and can be encoded in HRRR or other weather models as an alternative to the existing default lake depth configuration. The maximum depth in each lake in GLOBathy was not derived from observed depths (like GLDBv2), but instead calculated based on a random forest regression of various lake characteristics (for details and further discussion, see Khazaei et al., 2022; Zhan et al., 2023). The spatially heterogeneous depths were then calculated using a linear relationship between maximum depth and the distance to shore.

In order to incorporate GLOBathy into the model simulations, the original GLOBathy high-resolution data was resampled onto a 3 km × 3 km grid to accommodate the grid resolution of the existing HRRR-CLM-lake model. We experimented with multiple resampling techniques (including bi-linear, cubic, mean, and mode resampling) and visually compared the results to GLOBathy on its native resolution. Results from at least one previous lake study (Choulga et al., 2019) indicate that the mode is a suitable resampling method, however that study utilized a coarser (9 km × 9 km) grid. We found (results not shown) that resampling using the mode resulted in unrealistic spatial patterns, and ultimately determined that bi-linear resampling was the most suitable for our application. Additionally, bi-linear resampling minimized pixels that had depths shallower than 2 m which we viewed as a benefit to avoid the aforementioned "thin grid cell" problem. The resampled bathymetry (Figure 2) was encoded into the CLM framework using static geogrid input files. For the flatbottom case (GLDBv2), definition of the geogrid file is described in Benjamin et al. (2022).

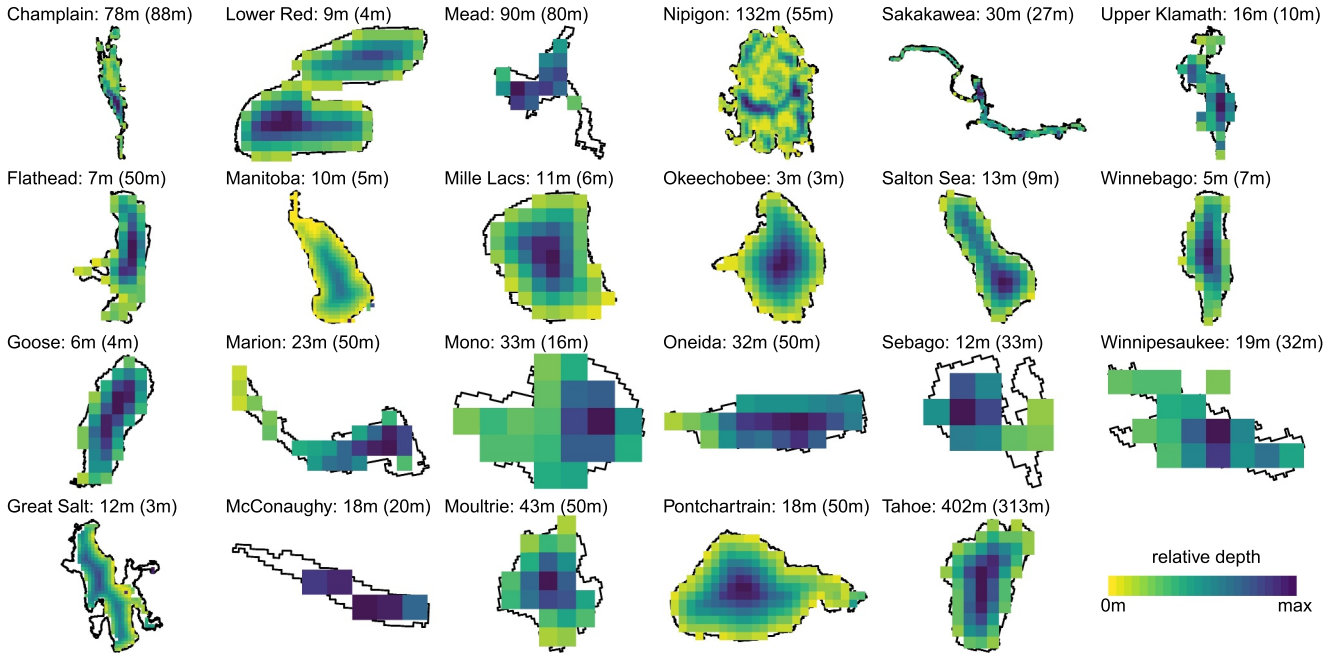

Figure 2. Graphical representation of the GLOBathy data set resampled onto the 3 km × 3 km grid in the CLM-lake model for each of the 23 lakes in our study. Coloring of grid cells is based on depth relative to a given lake's maximum depth in GLOBathy which is listed to the right of each lake name. The max depth from the original, flatbottom bathymetry is listed in parentheses.

2.4. Model Intercomparison and Validation

In the first step of our model analysis, we compared simulated lake-wide average (i.e., across all grid cells) lake surface temperatures (LST), ice cover, and evaporation from the flatbottom (default) configuration to those of the GLOBathy configuration. These variables were selected specifically for their role in regulating heat fluxes within each lake, and regulating both moisture and thermal fluxes between lake and atmosphere (Fujisaki-Manome et al., 2017; Liang et al., 1994; Xiang et al., 2017). Further, we believe that a model intercomparison (based on these three variables) can provide insight into the relative sensitivity of HRRR model output to different lake bathymetry configurations.

Occasionally some grid cells experienced erratic, non-physical lake surface temperatures for a few timesteps which was likely caused by the use of a large numerical integration timestep (3,600 s compared to 20 s for CLM-lake in the HRRR model). We filtered these values out using a generous acceptance window of −40°C to 50°C (water temperature below 0°C effectively indicates the surface temperature of ice.) Values outside this range were set to "NA" such that they were not included in any comparison or validation results. This filtering impacted six lakes in which filtered values accounted for less than 0.5% of the simulation period, with one exception, Lake Oneida, which had less than 3% of time window filtered.

In addition to lake-wide daily average surface temperatures, we also compared subsurface temperature evolution at many locations within the 23 lakes studied. Though subsurface temperatures are of second-order influence for numerical weather prediction, the presence and location of a thermocline is an important factor in modeled LST. Anderson et al. (2021) showed how changes in subsurface temperatures can influence the timing of stratification in a large lake (Michigan) which certainly impacts the vertical heat distribution and, in turn, surface conditions. In order to also address changes in lake dynamics at sub-daily time scales, we evaluated differences in diurnal temperature range (i.e., daily max minus daily min LST) based on the hourly model output.

We then assessed the relative skill of each model configuration (i.e., validation) by comparing simulated LST to observations. For satellite based validation, lake-wide average LST was calculated from the model output at 00 UTC to compare to the daily observations (Because the remote sensing data is collated from multiple instruments, there is no true "valid time" but we found that both simulations showed the best agreement with observations at 00

UTC). For in situ validation, we compared specific grid cells in the model to the co-located observations where available. Ice onset and breakup was also validated for the four lakes in our study that had reported dates.

Our validation provides what we believe is the first comprehensive assessment of lake physics representation in the HRRR-CLM-lake system, regardless of the bathymetry configuration. This assessment has the potential to help guide not only the prioritization and selection of a particular (potentially new) bathymetry data set, but also to evaluate the potential utility of using HRRR-CLM-lake output as a new continental-scale foundational data set that could provide better understanding of lake properties under changing climate as well as shifts in anthropogenic use and demand (Gronewold & Stow, 2014; Hampton, 2013; Yao et al., 2023).

2.5. Validation Data

Remote sensing lake surface temperature (LST) data was acquired from the European Center for Medium-Range Weather Forecasting's (ECMWF) Sentinel-based Copernicus Program Climate Data Store (Copernicus Climate Change Service (C3S), 2020). These data are stored as daily gridded NetCDF files at a 0.05° resolution and are based on aggregating information from a variety of satellite instruments including Advanced High Resolution Radiometers, Along Track Scanning Radiometers, the Sentinel-3 Sea and Land Surface Temperature Radiometer, and MODIS (Kilpatrick et al., 2015; Militino et al., 2020). To account for time periods when remote sensing imagery (and therefore LST data) is obscured by cloud cover, the Copernicus data was filtered using their pixel-based rating system in which each pixel is assigned a value from 1 to 5, where a value of 5 represents the highest quality data (for a related analysis, see Madonna et al., 2023). More specifically, on each day of the period of our study, we first filtered the Copernicus data to only include grid cells with a quality value of 3 or higher. We then identified, and removed from further analysis, any lakes for which less than 50% of the entire lake surface area data was valid (i.e., with a quality value greater than or equal to 3) on a given day. Although it resulted in a significant loss of validation data, this filtering process was important to ensure data quality. Nineteen of 23 lakes had available satellite data for validation and only a fraction of 2019 (on average 29%) was available for validation.

Data for our in situ validation was derived from a previous study (Sorensen et al., 2024) that includes surface and subsurface temperature data from 134 sites divided among 29 large lakes collected from a variety of sources. The temporal range of this data varies by lake and by site, and it spans approximately a decade in total, but nearly all lakes include data from 2019. For each in situ data point, we identified the grid cell in CLM-lake that contained that point and extracted the associated hourly time series of lake surface temperature. The spatio-temporal density of observations varied greatly from one observation location to another; as a result, not all lakes or seasons are equally represented by observational data.

3. Results

3.1. Overview

Our results are organized as follows: First, we identify how bathymetry was changed when applying GLOBathy on a lake-by-lake basis. Next, we focus on intercomparison by examining lake-wide differences in LST, ice cover, and evaporation for all lakes studied and subsurface temperature for a few specific locations as examples of where bathymetry change had noticeable impacts on thermal structure. We then validate LST at broad scales (lake-wide, daily) based on remote observations and relate the differences in LST performance back to changes in bathymetry. Lastly, we validate LST at finer scales for specific grid cells where in situ observations were available during 2019.

These results are summarized in Table 1, where the three column groups loosely correspond to the following three subsections: bathymetry comparison, model output intercomparison, and satellite validation (The in situ validation data was highly inconsistent in space and time so was not validated quantitatively). The first three columns show lake surface area based on the model domain, followed by the mean lake depth for each simulation. The middle three columns show, respectively, the deviation, the minimum difference and the maximum difference between daily averaged lake-wide LST in the two simulations. Although there is considerable variability from lake to lake, we note there is often symmetry about 0°C between the min and max LST departures (i.e., neither simulation had consistently lower or higher temperature extremes). The last two columns show the time fraction of valid satellite data and difference between model simulations (GLOBathy minus flatbottom) in root-mean-square error based on satellite data such that negative values indicate skill improvement when applying GLOBathy.

Table 1
Summary of Physical Characteristics and LST Statistics for Selected Lakes Over the 12-Month Period of 2019.

Name	Area (km^2)	$\overline{z}_{flat}$ (m)	$\overline{z}_{glob}$ (m)	*RMSD* (°C)	ΔT_{min} (°C)	ΔT_{max} (°C)	Sat. obs. (% year)	ΔRMSE (°C)
Champlain	1,089	42	28	0.7	−5.2	0.8	17	−0.01
Flathead	477	50	3	2.8	−6.6	7.8	26	1.38
Goose	342	4	3	0.4	−1.1	3	8	−0.3
Great Salt	2,817	3	4	0.3	−0.5	1.5	32	−0.11
Lower Red	1,143	4	4	0.8	−4.2	4.7	18	−0.17
Manitoba	2,907	5	2	1.1	−3.6	3.9	23	0.12
Marion	216	50	12	1	−4.8	2.2	25	−0.71
McConaughy	45	20	16	1	−3.8	3.4	0	NA
Mead	126	80	58	0.4	−0.8	0.9	0	NA
Mille Lacs	513	6	4	0.7	−2.9	3.2	18	−0.11
Mono	180	16	16	0.5	−1.4	1.2	56	−0.25
Moultrie	189	50	19	0.8	−4.3	1.6	30	−0.57
Nipigon	5,229	55	37	1.4	−4.4	8.5	19	0.06
Okeechobee	1,359	3	1	0.6	−2.2	2	37	0.07
Oneida	225	50	19	0.6	−2.3	2.6	20	−0.44
Pontchartrain	1,674	50	7	1.5	−5	3.3	0	NA
Sakakawea	900	27	13	2.2	−10.7	5.7	15	−0.33
Salton Sea	891	9	5	0.5	−1.8	1.3	59	0.11
Sebago	90	33	7	1.9	−5.3	4.9	23	0.25
Tahoe	477	313	190	0.7	−1.3	1.7	51	−0.39
Upper Klamath	288	10	8	0.3	−1.3	1.6	31	−0.09
Winnebago	495	7	2	1.4	−3.2	5.3	21	−0.21
Winnipesaukee	153	32	9	2	−8.2	4.9	0	NA
mean	949	40	20	1	−3.7	3.3	28	−0.09

Note. Columns 1–3 show lake surface area as well as mean depth for flatbottom and GLOBathy. Columns 4–6 show differences in lake-wide daily mean LST: root mean square deviation (through time) between model simulations, followed by min and max difference (GLOBathy-flatbottom) on any given day. The rightmost two columns address satellite validation: availability of valid satellite obs. throughout 2019 and the difference (GLOBathy-flatbottom) in root-mean-square error between model and satellite obs. A negative value means lower error for GLOBathy than flatbottom experiments.

3.2. Bathymetry Comparison

Our analysis of bathymetry changes in 23 lakes found that, compared to the default configuration, mean lake depth using the GLOBathy configuration either decreased (18 lakes), was unchanged (three lakes) or had marginal increase (one lake). This finding has important implications for the distribution of heat content within each lake grid and the propagation of heat content into changes in surface temperature, ice cover and other variables.

The most profound decreases in mean lake depth (Table 1, columns 2 and 3) include Flathead (94% reduction or 47 m decrease), Pontchartrain (86%, 43 m) and Sebago (80%, 26 m). The average depth decrease among 23 lakes was 50% (from 40 to 20 m). As described in Section 3.3 of Benjamin et al. (2022), unknown lake depths in GLDBv2 (and in HRRR-CLM-lake) were assigned an often too-deep value of 50 m (evident for five lakes in Table 1) even if later corrected in GLDBv3 to a more reasonable value of 10 m (Choulga et al., 2019). Although the mean depth decreased in the GLOBathy configuration in almost all lakes in our study, the maximum depth *increased* for approximately half of the lakes (see Figure 2) due to the nature of the spatial variation in depth within the GLOBathy algorithm (see Khazaei et al., 2022, for more details). Though each lake is shown at the same relative scale in Figure 2, lake surface area varied considerably which can be seen by carefully looking at the 3-km grid-cell width. Note that larger lakes (e.g., Champlain, Manitoba, Nipigon) are fairly well resolved horizontally, whereas smaller lakes (McConaughy, Sebago) are not. The impact of the changes in bathymetry can be assessed throughout our results in the following two subsections.

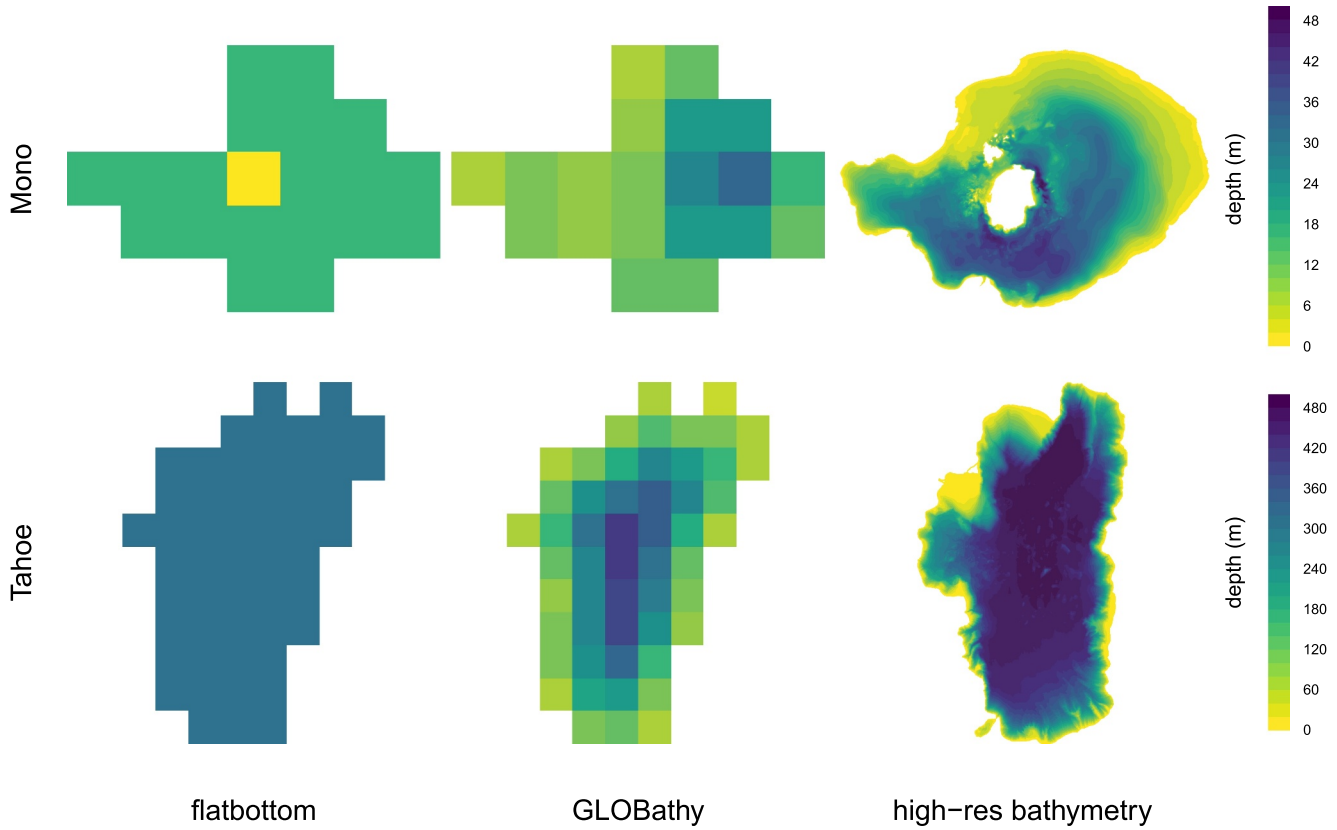

Figure 3. Comparison of flatbottom versus GLOBathy versus high resolution bathymetry for Mono Lake (Raumann et al., 2002) and Lake Tahoe (Gardner, 1999).

It's important to note that even if a lake's mean depth did not change (or change significantly) using GLOBathy, the spatial distribution almost certainly did change (see Figure 2) and the modeled LST was impacted. For example, Mono Lake had virtually no change in mean depth but exhibited a 0.5°C deviation between simulations. The largest daily, lake-wide LST departures were exhibited for Sakakawea (≈10°C) which experienced a 44% reduction in mean depth in GLOBathy.

Figure 3 compares the bathymetry in our CLM-lake simulations to real-world, high resolution bathymetry. For Mono Lake, the max depth in flatbottom is grossly underestimated (16 vs. 48 m in reality). In GLOBathy, max depth is doubled (33 m) although the location of max depth is not accurate compared to the real-world data, and the island is no longer represented. Lake Tahoe is an interesting case since, in reality, it is a fairly flat-bottomed lake with near uniform depth. The true average depth is around 300 m which is much better represented by flatbottom than GLOBathy (see Table 1). However, GLOBathy does a better job of capturing the bathymetry in near-shore areas whereas the shoreline in the flatbottom case is represented as a 300 m cliff from shore to lakebed.

3.3. Model Intercomparison

The seasonal evolution of lake-wide LST differences showed large variations throughout the year (see Figure 4). As a lake becomes fully covered with ice, LST has diminished importance for atmospheric processes so hatched lines (superimposed on the temperature difference graphic) indicate days when ice cover was greater than 90% in both simulations. As expected, the decrease in bathymetry from flatbottom to GLOBathy reduced total heat capacity of the lakes and, in general, lakes cooled faster in the fall and warmed faster in the spring. There's also an unexplained relative warming in the GLOBathy simulation during mid-January through March for some higher latitude lakes (Nipigon, Sakakawea, Sebago, Winnebago, Winnipesaukee). In all cases however, these lakes are mostly covered with ice and therefore thermally insulated from the atmosphere.

It's worth noting that several lakes (Champlain, Great Salt, Upper Klamath) show very little differences in LST between simulations. This is likely due to small changes in bathymetry (Great Salt and Upper Klamath) or in the

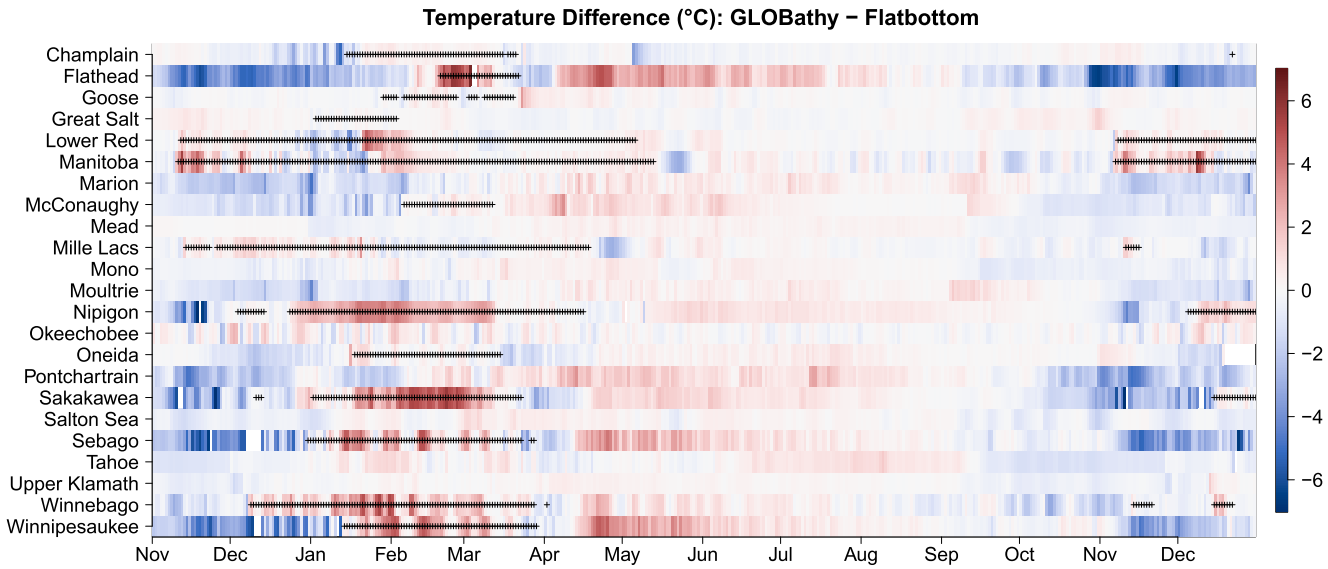

Figure 4. Difference between simulated lake-wide daily average LST (GLOBathy—flatbottom) from November 2018 through December 2019 (colorscale is constrained to show variation; see Table 1 for extremum.) Hatched lines indicate periods when both simulations had ice cover exceeding 90%.

special case of Champlain, the fact that default bathymetry was based on real data and not spatially uniform (i.e., Champlain was not actually flat-bottomed in the default case).

In general there was a considerable increase in lake ice cover when using the GLOBathy data set (see Figure 5). This may be cause for concern as previous work found CLM-lake (with default bathymetry) to overestimate ice cover in larger lakes (C. Xiao et al., 2016). The range of color (compared to colorbar limits) illustrate there are periods of time when GLOBathy has nearly 100% ice cover, but flatbottom has none; conversely, flatbottom is at most 30% higher than GLOBathy and these periods were short-lived. The increased ice cover in GLOBathy tends to occur primarily during the fall, presumably due to shallower grid cells cooling faster using GLOBathy which is consistent with the LST differences. Furthermore, the timing of ice onset is shifted more than that of ice breakup using GLOBathy (i.e., wider blue regions during fall vs. spring in Figure 5.) This may indicate that bathymetry and lake thermodynamics are more important during modeled ice formation whereas atmospheric conditions (identical for both simulations) dictate ice breakup.

Naturally, there is strong negative correlation between LST and ice cover differences. Though the GLOBathy simulations generally warm faster in the spring, the ice cover appears to persist longer for several lakes. This may be due to higher ice concentrations, since ice has a high albedo and acts as an insulator between lake and atmosphere. Goose and Tahoe both seem to defy the general LST difference trend and also experience primarily

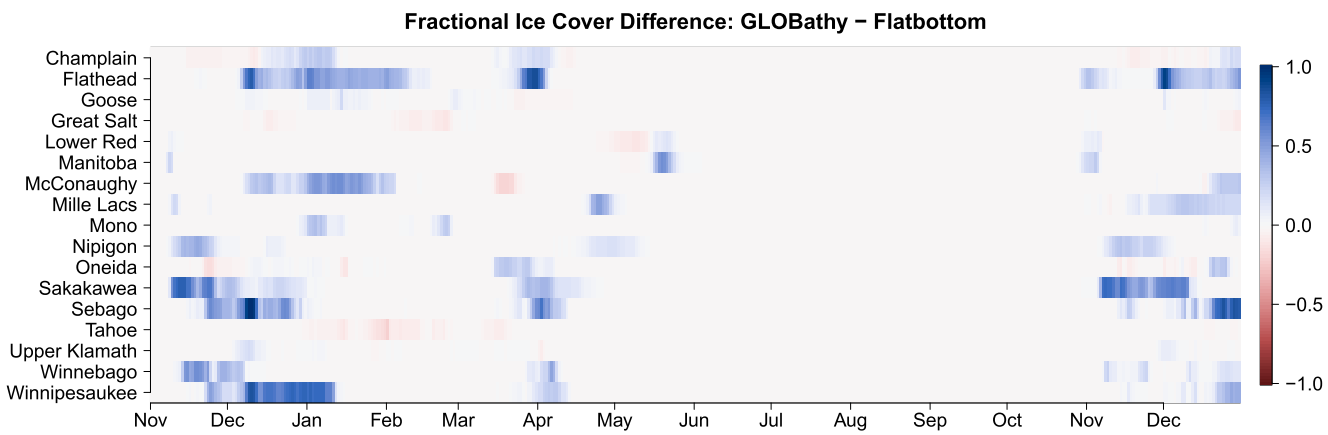

Figure 5. Difference in ice cover between model simulations for the 17 lakes (of 23 studied) that experienced any freezing throughout the simulation.

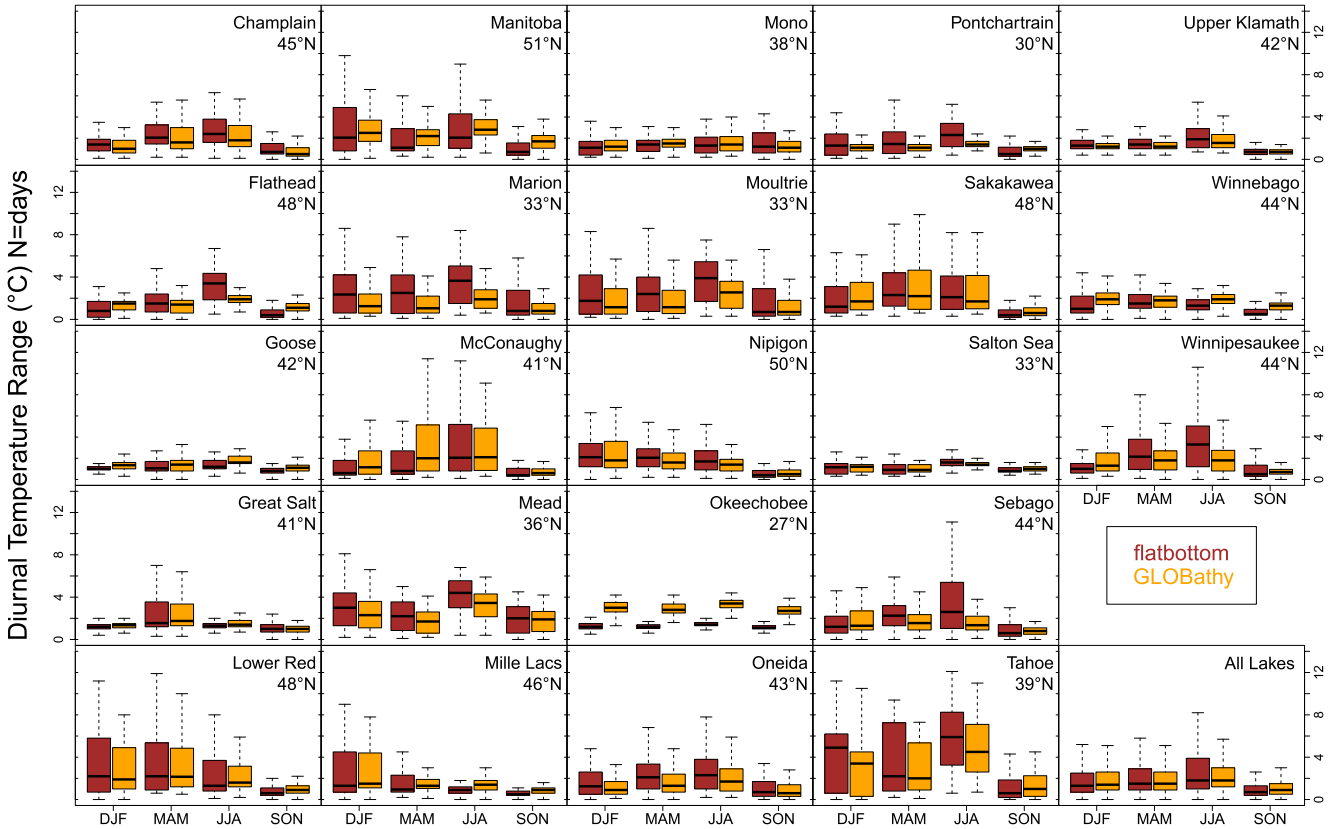

Figure 6. Lake-wide diurnal surface temperature range (i.e., $T_{max} - T_{min}$ based on hourly temperature output) categorized by meteorological season and lake. The members of each distribution are temperature ranges on all days in a given season such that N≈90. Outliers (not shown) were typically <20°C.

less ice cover using GLOBathy. Goose Lake's mean bathymetry changed very slightly between simulations; Tahoe had a large change, but under both bathymetry configurations it is an order of magnitude deeper than other lakes in terms of max depth (Figure 2) and mean depth (Table 1) so it may behave differently from other lakes when bathymetry is changed. Some lakes (Nipigon, Winnebago, Sakakawea) exhibited significantly higher LST during winter while showing negligible difference in ice cover, which is most likely due to water (or ice) temperature values well below zero (see discussion of Figure 9). The fact that LST represents the ice temperature further warrants the need for masking out LST differences during periods of high ice cover in Figure 4.

Daily LST ranges were organized by lake, season, and simulation case and then analyzed in order to compare the influence of bathymetry on diurnal variations (see Figure 6). Some interesting seasonal patterns can be noted before considering the comparison between simulations. For example, higher latitude lakes tend to have larger diurnal temperature variation than lower latitude lakes. This is particularly true during winter and summer for some of the northernmost lakes (Lower Red, Manitoba, Mille Lacs, Sakakawea)—likely due to large air (and air-water) temperature differences during warm and cool seasons. The distribution of diurnal LST range for all lakes (lower right panel) show very little difference between simulations, but several individual lakes do show a significant difference in GLOBathy. Most notably, Okeechobee's median LST range approximately doubled for all seasons which is not unexpected considering that mean depth decreased from 3 to 1 m using GLOBathy. Conversely, Marion and Moultrie both had a relative reduction in mean bathymetry similar to Okeechobee (see Table 1) but experienced *decreased* diurnal variation year-round. This contradiction is difficult to reconcile, but it may be related to the larger absolute mean depths of Marion and Moultrie, both of which remained >10 m in GLOBathy.

Evaporation was integrated for each lake and season and the results from each simulation case were compared (see Figure 7). The seasonal distribution of total evaporation varied considerably from lake to lake. Overlake evaporation was smallest for most higher latitude lakes, particularly during winter and spring ("DJF," "MAM"),

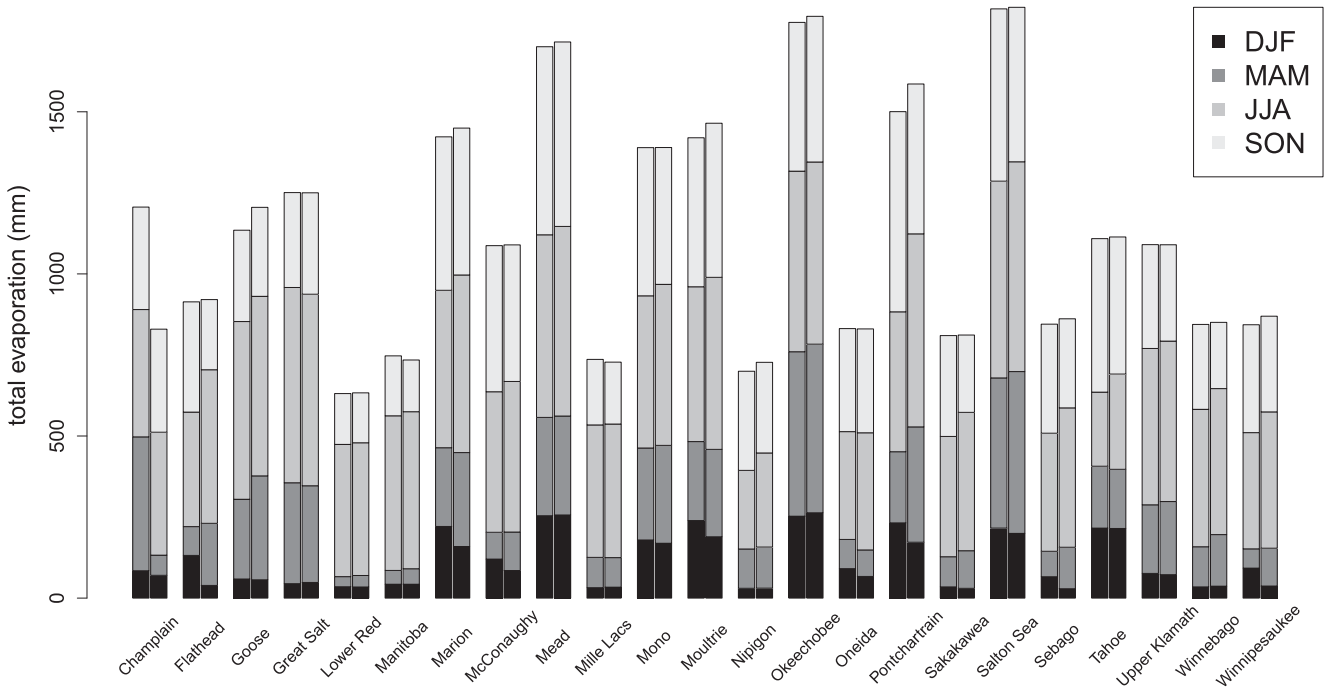

Figure 7. Overlake evaporation for flatbottom (left bars) and GLOBathy (right bar), integrated from hourly output over each meteorological season.

likely due to low LST (and therefore small air-water temperature differentials) and the presence of ice cover, which has been suggested to inhibit evaporation and thus modify the timing of evaporation during fall, winter, and spring (Lenters et al., 2013). Conversely, lower-latitude lakes generally experienced the highest total evaporation, which was more evenly distributed year-round. Despite large (short-lived) departures in surface temperature between simulations, total evaporation was largely the same in each simulation. The most notable exception was Champlain, which had considerable reduction in evaporation in the GLOBathy case that occurred mostly during the spring ("MAM"), despite having fairly consistent temperature and ice cover between simulations during this period (see Figures 4 and 5). Several lakes (Flathead, McConaughy, Sebago, Winnipesaukee) that experienced higher ice cover using GLOBathy had decreased winter evaporation, which was offset by greater spring evaporation—likely due to faster spring warming (see Figure 4) despite the increased ice. Some lower latitude lakes that didn't freeze (Marion, Moultrie) also exhibited lower evaporation during winter (presumably due to decreased LST), but had less pronounced "make up" evaporation the following spring. Pontchartrain also experienced decreased evaporation during winter but had greater total evaporation using GLOBathy due to increased evaporation in other seasons. Comparison of hourly evaporation rates between simulations (see Supporting Information S1) revealed small departures (typically ~0.1 mm, outliers ~10 mm) and no consistent bias. This finding suggests that bathymetry's main impact was on the timing of evaporation rather than the total magnitude throughout the year as can also be seen in Figure 7.

Temperature profile comparisons reveal how bathymetry changes influenced thermal distribution in the water column. Figure 8 shows subsurface temperature for four individual grid cells. The top (bottom) two panels show examples where bathymetry increased (decreased) using GLOBathy. The left two panels show a qualitative change in lake temperature evolution in regard to thermal stratification. In Lower Red, a thermocline never developed in this particular grid cell for flatbottom, whereas some moderate stratification occurred for several months (mid-June through September) for the GLOBathy case. Upper Klamath shows a similar but inverted change where a thermocline was present in flatbottom but not in GLOBathy. The right two panels show cases where qualitative changes in stratification were not seen despite large changes in depth. For Great Salt Lake, depth was approximately doubled in GLOBathy, but a thermocline never developed in either case. For Marion, lake depth decreased by a factor of 5 and though the vertical heat distribution changed, the presence and timing of stratification was not impacted considerably. This may be due to the comparatively (still) deep grid cell in Marion. Although these four locations may imply that 10 m is a threshold for stratification to occur in CLM-lake, this is not

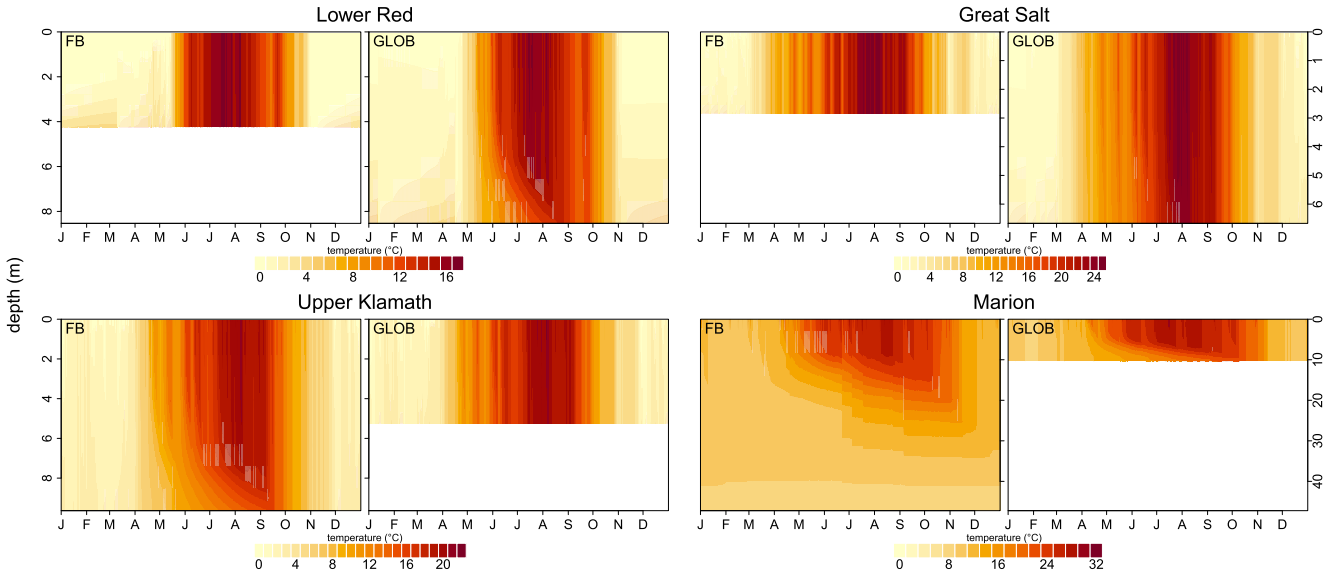

Figure 8. Water column temperature versus time at four locations for flatbottom (left subpanels) and GLOBathy (right subpanels).

consistent at other locations (see Supporting Information S1) and so other factors, such as latitude, likely contribute to whether or not a given lake grid cell stratifies.

3.4. Remote Validation

Satellite data validation was quantified via RMSE through time in Table 1, but additional insight can be gained by visualizing the simulated and observed temperatures as time series (see Figure 9). Due to the large seasonal range of temperatures (i.e., broad *y*-axis limits), it's difficult at times to see differences between simulations, but there are some points in time when simulations deviate considerably. During winter months, many lakes exhibit surface temperatures well below zero (effectively representing the surface temperature of ice), which are not shown in order emphasize the observed temperature data (always non-negative) and also constrain the *y*-axis to show variation. The seasonal distribution of remote validation data can also be visualized–Champlain, for example, had reasonably consistent observation data year round despite having only 17% of days in 2019 available (see Table 1). Although the improvement is subtle in Figure 9, skill is visibly better in GLOBathy for the three lakes that showed the greatest root-mean-square error decrease (Marion, Moultrie, Oneida) in Table 1. For seemingly the same reason (shallow water ⇒ less thermal mass), GLOBathy performs worse for other lakes like Sebago and Flathead, where summer temps are too high and winter temps are too low. Most lakes show little differences between simulated LST relative to the error between model and observations particularly during summer where both simulations are too cold. This is particularly true for Lower Red, Sakakawea, and Mille Lacs where summer biases (model vs. obs) are on the order of 10°C and the difference between simulations is on the order of 1°C or less. In other words, the GLOBathy simulations did not mitigate the biases present in the flatbottom simulations.

The characteristics of each large lake varied considerably, as did the relative change in bathymetry when applying GLOBathy. In order to address model sensitivity to these characteristics we related changes in model LST differences (Figure 10 left panel) and model LST performance (Figure 10 right panel) to the change in mean model depth (Figure 10 is a graphical representation of Table 1 where the horizontal axes are both relative difference between columns 2 and 3 and the vertical axes are columns 4 and 8, respectively). Recall that the deviations are reported based on spatial (lake-wide) and temporal (24 hr) averages, such that small departures have significant implications. The results are indicative of a strong sensitivity to lake depth—11 out of 23 lakes have LST deviations greater than or equal to 1°C and three lakes have deviations exceeding 2°C. These changes are comparable to those observed due to climate change (Austin & Colman, 2007; O'Reilly et al., 2015), which underscores the importance of bathymetry selection since climate projections are often based on similar earth system models.

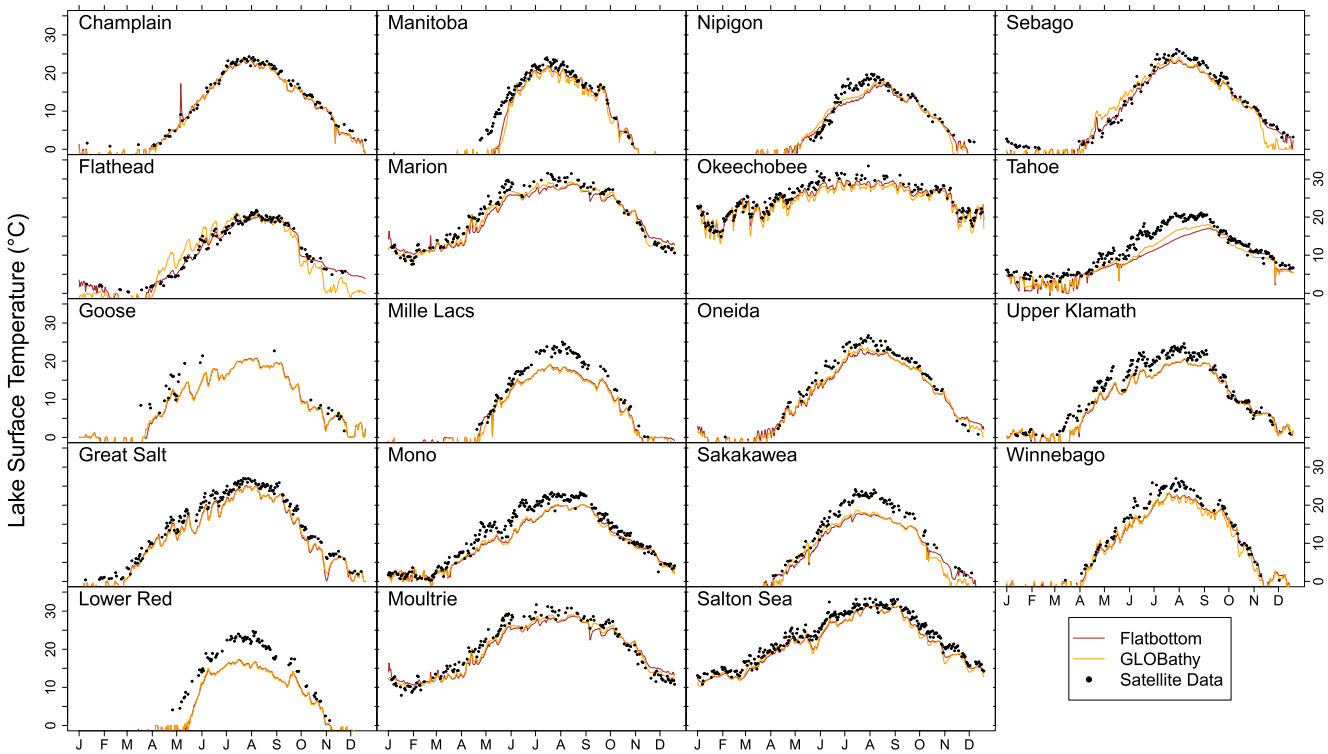

Figure 9. Comparison between model simulations and remote observations for lake-wide average LST for the 19 lakes (of 23 studied) with available validation data. Note that *y*-axis limits are constrained to show only positive temperature values (Subzero values were present in the model simulations and represent surface temperature of ice).

The relationship between mean depth reduction and RMSD appears somewhat linear with larger changes causing larger deviations in temperature. As stated previously, the mean depth change does not reflect the spatial changes in bathymetry. For example, Lower Red and Mono show non-zero RMSD's despite having no change in mean bathymetry. There does not appear to be a strong relationship between either RMSD or RMSE and lake surface area or latitude. Note, however, that most of the lower latitude lakes had smaller RMSD, per change in depth, (darker colored dots in the left panel of Figure 10 fall below the $y = x$ line), implying that these southern lakes were perhaps less sensitive to bathymetry changes (This relationship is based on a small subset of lakes and is likely complicated by the absence vs. presence of ice cover in southern vs. northern lakes, so more work is needed to determine if this is broadly true in CLM-Lake). Unlike RMSD, model *performance* does not appear to have a direct relationship with depth change (i.e., changing the depth more does not consistently result in more accurate surface temperatures). The RMSE difference varies a great deal from lake to lake (Table 1 and Figure 10) with extreme values of 0.74°C improvement in GLOBathy (Marion) to 1.38°C performance decrease (Flathead). On average, RMSE skill increases only 0.09°C in GLOBathy. Even for the lakes that showed the largest decreases in RMSE, the improvement was only a fraction of the error between model and observations which ranged between 1.3 and 6.0°C (median 2.7°C.) In short, the improvement in model performance by changing bathymetry was small compared to the total error between (either) simulation and observations.

Ice validation data was quite limited, but four lakes had at least one observation for either ice on or ice off from Sharma et al. (2022) (as can be seen in Figure 11). Observed ice on definition varied and was described as "completely ice covered" or "no commercial boat traffic," while observed ice off was "Mostly free of ice and navigable by boat." Therefore, in our simulations we defined ice onset as the first instance of ice >90% and ice off as the last instance of ice ≥10%. These values are somewhat arbitrary, but we note that the on/off date was fairly insensitive to the threshold choice. For example, using 70% and 30% instead typically resulted in the on/off date shifting less than 1 week. This illustrates that modeled lake ice cover is somewhat bimodal in the sense that the lakes quickly transition from open water to near full ice coverage, which can be confirmed by lake-wide ice time series plots (not shown).

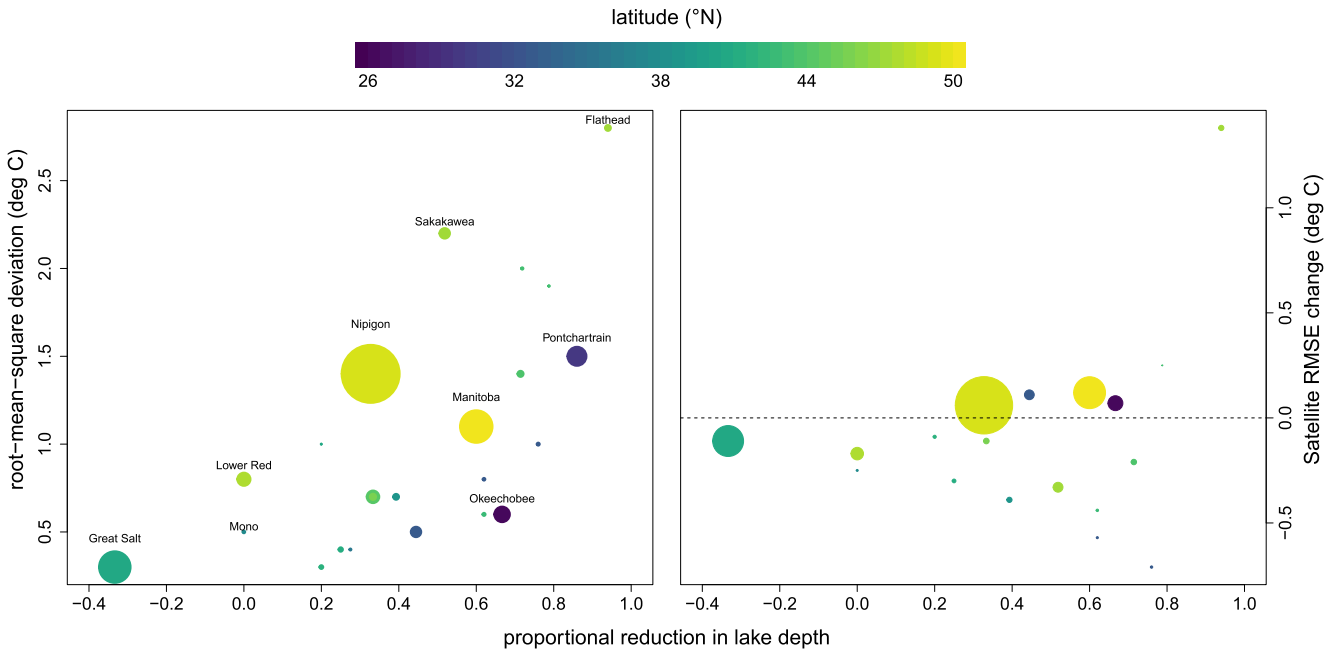

Figure 10. Relationship between proportional mean lake depth reduction (from flatbottom to GLOBathy) and the associated lake-wide LST RMSD between simulations (left) and LST skill difference: GLOBathy—flatbottom, validated by satellites (right). Point size is relative to lake surface area and color indicates latitude.

3.5. In Situ Validation

In situ data was used to qualitatively validate specific grid cells in both simulation cases (see Figure 12). At most locations, the overall seasonal pattern remains consistent between the flatbottom and GLOBathy simulations. However, the two model runs sometimes exhibit different timing or magnitude of daily fluctuations, such as at the third and fourth sensor locations on Lake Oneida (ONE03, ONE04), the latter of which also experiences approximately 10°C cooler temperatures during the summer months in the Flatbottom case than in GLOBathy. Across lakes and sensor locations, the two model runs are generally more similar to one another than to the observations. Similar to the remote (lake-wide) validation results, the in situ data were typically warmer during the summer than either model represented, particularly at CLE03.

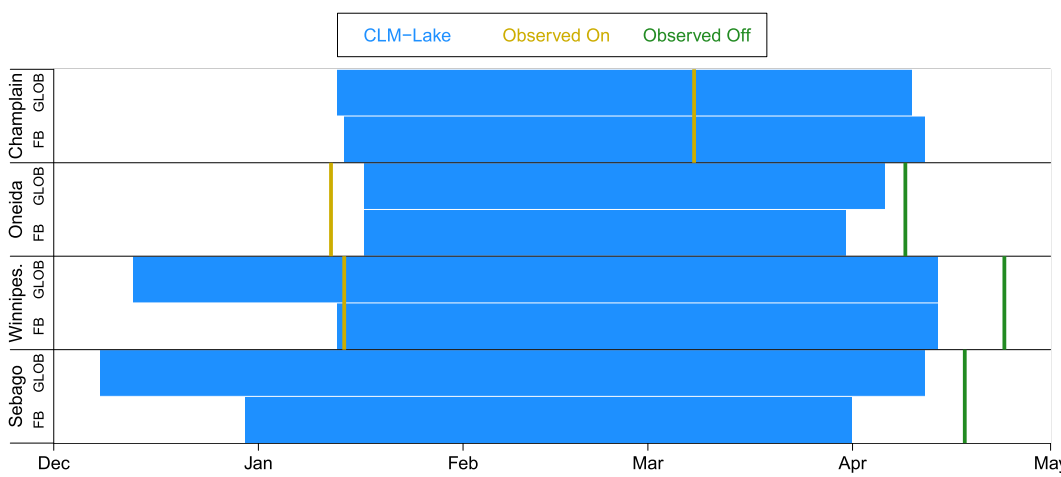

Figure 11. Ice on/off dates validation for four lakes which had at least one observation during 2019. Ice onset (breakup) is defined as the first (last) date that fractional ice cover exceeds 90% (10%) fractional cover.

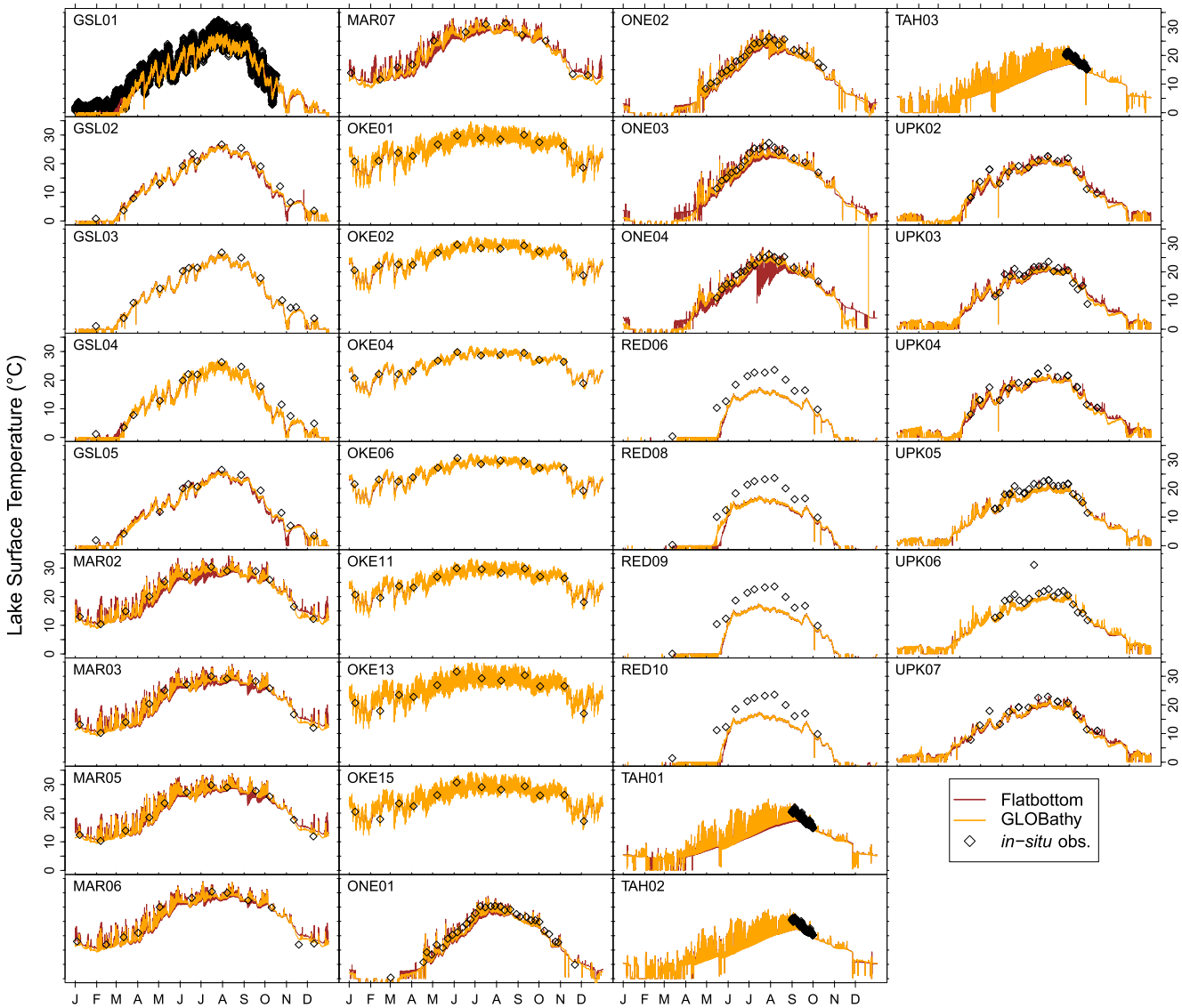

Figure 12. Lake surface temperature validation for all locations with 12 or more in situ observations in 2019. The *y*-axis is constrained to show only positive temperature values.

4. Discussion

Throughout the previous section, considerable differences in simulated lake surface conditions were identified as a result of modifying bathymetry. We now focus on the implications of these changes for 1-D lake models and numerical weather prediction systems. Although the simulation domain was restricted to lakes and primary results were focused on the lake surface, the potential impacts into weather models should be emphasized, since these models, in practice, are coupled with 1-D models like CLM-lake.

We recognize that running an *offline* version of the HRRR-CLM-lake system constrains the range of variables available for model intercomparison and validation to those directly associated with lake physical properties (such as LST and ice cover), rather than variables describing atmospheric conditions (and, potentially, more directly linked to meteorological phenomena). Nonetheless, we believe that a model intercomparison study that focuses specifically on relative impacts on lake physical processes has value to both the NWP research and operational forecasting communities. Put differently, we believe that it is reasonable, within the scope of this study, to presuppose that improvements (if any) to the simulation of lake physics properties in HRRR-CLM-lake

associated with a new lake bathymetry data set are very likely to propagate into improvements in simulated lake-atmosphere fluxes and other two-way interactions (Bryan et al., 2015).

As an example of lake-atmosphere sensitivity, Wright et al. (2013) showed that a difference of 3°C in LST in the Laurentian Great Lakes results in a significant change in both intensity and band location of lake effect precipitation using the same atmospheric model (Weather Research and Forecast) applied in the HRRR system. Our results showed that for 16 of 23 large lakes studied, temperature differences (averaged over an entire day and entire lake surface) were greater than 3°C. Although the largest lake studied (Nipigon) has approximately one fifth the surface area of the smallest Great Lake, the difference in bathymetry may have the potential to influence lake effect precipitation in the HRRR weather prediction system. Further work involving two-way coupled HRRR simulations will be critical to determine the magnitude of this influence.

It should be noted that although GLOBathy employs a more sophisticated algorithm for determining lake bathymetry, it may be less realistic than flatbottom bathymetry for some lakes (see Section 3.2 for examples). The flatbottom bathymetry may be a reasonable approximation of the true bathymetry for lakes which have near uniform depth and bathymetry selection may be somewhat irrelevant for smaller lakes whose shoreline could not be resolved accurately by any 3 km grid. For larger lakes with more complex bathymetry, the flatbottom assumption may be inappropriate and contribute to model error. It's also possible that due to the relatively simple representation of physics in 1-D lake models, the most realistic bathymetry data set does not result in the most accurate lake dynamics and lake surface representation. Using GLOBathy resulted in more realistic bathymetry for unknown-depth lakes (which were previously assigned a too deep value of 50 m), but simultaneously distorted grid cells nearshore (too shallow relative to surface area). At these shallow locations, the 1-D lake model—lacking a mechanism to represent horizontal transport, responded too quickly to atmospheric forcing due to decreased thermal mass.

The extent to which a particular selection of lake bathymetry leads to an *improvement* in HRRR-CLM-lake skill is less certain. Our study indicates that a heterogeneous bathymetry tended to lead toward a shallower mean lake depth and faster water temperature transitions in fall and spring. Understanding how these changes in surface conditions propagate into weather prediction systems requires more research; results from two-way coupled simulations may help elucidate this issue. Another potential solution is the implementation of 3-D lake models, perhaps on a case-by-case basis, in a manner similar to what has been implemented on the Great Lakes. Clearly, the time and computational investment in a 3-D model for the Great Lakes is high, but proportional to the impact of the Great Lakes on the climate and water resources of central North America. An investment in 3-D models for smaller, but also important (to regional weather, climate, and water resources) lakes, should be considered as a next step following the research presented here.

5. Summary

The selection of lake bathymetry in a 1-D model can have a profound impact on lake surface conditions. In our study, we found that bathymetry changes led to simulated differences in lake surface temperature, ice cover, evaporation and sub-surface thermal structure. This sensitivity to bathymetric configuration within operational weather prediction systems should be reconciled.

In order to explore said sensitivity, we conducted parallel simulations of the 1-D lake model used by NOAA's operational weather model and then evaluated differences between simulations and used both in situ and remote observations to evaluate model performance. The first simulation used the default, or "flatbottom," bathymetry currently implemented in the existing operational model. The second simulation used spatially varying bathymetry, known as "GLOBathy" which was developed using an algorithm based on known lake properties. Applying GLOBathy to lakes in our study, on average, reduced lake-wide bathymetry by a factor of 2 when compared to the GLDBv2 flatbottom (from 40 to 20 m). Another bathymetry option (not explored here) is use of GLDBv3 with 10-m depth for unknown lakes (instead of 50 m) as suggested by Benjamin et al. (2022) following Choulga et al. (2019).

The decrease in mean bathymetry led to reduced thermal mass in the GLOBathy simulation which exhibited faster seasonal transitions and a longer ice season (predominantly due to earlier ice onset). Diurnal temperature cycle was impacted considerably in some lakes and some seasons but showed no consistent trend throughout the 23 lakes studied. Total (year-round) evaporation was largely unchanged despite some moderate seasonal shifts in

some lakes. Daily, lake-wide temperature departures between simulations were as great as 10°C (mean throughout lakes 3°C) with daily root-mean-square deviations greater than 2°C (mean 1°C).

The lake surface temperature skill between simulations showed marginal difference in most cases with 13 out of 19 lakes validated showing a change in RMSE of less than 0.25°C in magnitude. Six lakes showed improved skill in GLOBathy (ΔRMSE < -0.25°C) and one lake showed decreased skill (ΔRMSE > 0.25°C). Because the median RMSE for all simulations was more than 2°C, it could be argued that no lakes showed significant improvement. However, the intercomparison is equally, if not more, important than the model validation. Neither bathymetry data set is perfect, and it is likely that neither is best suited for CLM-lake, but the broader implication is that lake thermodynamics in this context are heavily influenced by bathymetry selection. Since accurate lake surface conditions are critical for numerical weather prediction skill near large lakes, more work is needed to determine the most ideal bathymetry for this application. As the atmospheric component of NWP evolves, the representation of lakes must evolve too; rethinking bathymetry selection is a logical starting point for improvement.

Data Availability Statement

The data (CSV and netCDF) and software (R) used for analyses in this study are available via Zenodo and GitHub at https://doi.org/10.5281/zenodo.13953004. The majority of intercomparison between model simulations and skill assessment versus observations were performed on the lake-wide temperature, ice, and evaporation data which are available as CSV files. Additionally, spatially distributed depth, temperature, and ice files are available as netCDF.

Acknowledgments
Funding for this study was provided through the NOAA JTTI program (UM-CIGLR Grant NA21OAR4590177, NCAR Grant NA21OAR4590178, and UC-CIRES Grant NA21OAR4590179). Support was also provided by NOAA GLERL and NOAA ESRL. ADG was partially supported through funding from the National Science Foundation Global Center's Program (NSF award No. 2330317). This research was supported in part through computational resources and services provided by Advanced Research Computing (ARC), a division of Information and Technology Services (ITS) at the University of Michigan, Ann Arbor. This is GLERL contribution number 2063.

References

Anderson, E. J., Fujisaki-Manome, A., Kessler, J., Lang, G. A., Chu, P. Y., Kelley, J. G. W., et al. (2018). Ice forecasting in the next-generation great lakes operational forecast system (GLOFS). *Journal of Marine Science and Engineering*, *6*(4), 123. https://doi.org/10.3390/jmse6040123

Anderson, E. J., Stow, C. A., Gronewold, A. D., Mason, L. A., McCormick, M. J., Qian, S. S., et al. (2021). Seasonal overturn and stratification changes drive deep-water warming in one of Earth's largest lakes. *Nature Communications*, *12*(1), 1688. https://doi.org/10.1038/s41467-021-21971-1

Austin, J. A., & Colman, S. M. (2007). Lake superior summer water temperatures are increasing more rapidly than regional air temperatures: A positive ice-albedo feedback. *Geophysical Research Letters*, *34*(6), L06604. https://doi.org/10.1029/2006gl029021

Benjamin, S. G., Smirnova, T. G., James, E. P., Anderson, E. J., Fujisaki-Manome, A., Kelley, J. G. W., et al. (2022). Inland lake temperature initialization via coupled cycling with atmospheric data assimilation. *Geoscientific Model Development*, *15*(17), 6659–6676. https://doi.org/10.5194/gmd-15-6659-2022

Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., et al. (2016). A North American hourly assimilation and model forecast cycle: The rapid refresh. *Monthly Weather Review*, *144*(4), 1669–1694. https://doi.org/10.1175/mwr-d-15-0242.1

Briley, L. J., Rood, R. B., & Notaro, M. (2021). Large lakes in climate models: A great lakes case study on the usability of CMIP5. *Journal of Great Lakes Research*, *47*(2), 405–418. https://doi.org/10.1016/j.jglr.2021.01.010

Bryan, A. M., Steiner, A. L., & Posselt, D. J. (2015). Regional modeling of surface-atmosphere interactions and their impact on Great Lakes hydroclimate. *Journal of Geophysical Research: Atmospheres*, *120*(3), 1044–1064. https://doi.org/10.1002/2014jd022316

Choulga, M., Kourzeneva, E., Balsamo, G., Boussetta, S., & Wedi, N. (2019). Upgraded global mapping information for earth system modelling: An application to surface water depth at the ECMWF. *Hydrology and Earth System Sciences*, *23*(10), 4051–4076. https://doi.org/10.5194/hess-23-4051-2019

Copernicus Climate Change Service (C3S). (2020). Lake surface water temperature from 1995 to present derived from satellite observation. https://doi.org/10.24381/cds.d36187ac

Dowell, D. C., Alexander, C. R., James, E. P., Weygandt, S. S., Benjamin, S. G., Manikin, G. S., et al. (2022). The High-Resolution Rapid Refresh (HRRR): An hourly updating convection-allowing forecast model. Part I: Motivation and system description. *Weather and Forecasting*, *37*(8), 1371–1395. https://doi.org/10.1175/waf-d-21-0151.1

Fujisaki-Manome, A., Fitzpatrick, L., Gronewold, A. D., Anderson, E. J., Lofgren, B. M., Spence, C., et al. (2017). Turbulent heat fluxes during an extreme lake effect snow event. *Journal of Hydrometeorology*, *18*(2), 3145–3163. https://doi.org/10.1175/jhm-d-17-0062.1

Fujisaki-Manome, A., Mann, G. E., Anderson, E. J., Chu, P. Y., Fitzpatrick, L. E., Benjamin, S. G., et al. (2020). Improvements to lake-effect snow forecasts using a one-way air–lake model coupling approach. *Journal of Hydrometeorology*, *21*(12), 2813–2828. https://doi.org/10.1175/jhm-d-20-0079.1

Gardner, J. (1999). *Lake Tahoe marine geology and geophysics from field activity:i-1-98-lt* (Tech. Rep.). U.S. Geological Survey, Coastal Marine Geology. Retrieved from https://pubs.usgs.gov/dds/dds-55/pacmaps/lt_data.htm

Garnaud, C., MacKay, M., & Fortin, V. (2022). A one-dimensional lake model in ECCC's land surface prediction system. *Journal of Advances in Modeling Earth Systems*, *14*(2). https://doi.org/10.1029/2021ms002861

Gowan, T. A., Horel, J. D., Jacques, A. A., & Kovac, A. (2022). Using cloud computing to analyze model output archived in ZARR format. *Journal of Atmospheric and Oceanic Technology*, *39*(4), 449–462. https://doi.org/10.1175/jtech-d-21-0106.1

Gronewold, A. D., & Stow, C. A. (2014). Water loss from the great lakes. *Science*, *343*(6175), 1084–1085. https://doi.org/10.1126/science.1249978

Gu, H., Jin, J., Wu, Y., Ek, M. B., & Subin, Z. M. (2013). Calibration and validation of lake surface temperature simulations with the coupled WRF-lake model. *Climatic Change*, *129*(3–4), 471–483. https://doi.org/10.1007/s10584-013-0978-y

Hampton, S. E. (2013). Understanding lakes near and far. *Science*, *342*(6160), 815–816. https://doi.org/10.1126/science.1244732

Hostetler, S. W., & Bartlein, P. J. (1990). Simulation of lake evaporation with application to modeling lake level variations of Harney-Malheur Lake, Oregon. *Water Resources Research*, *26*(10), 2603–2612. https://doi.org/10.1029/wr026i010p02603

Khazaei, B., Read, L. K., Casali, M., Sampson, K., & Yates, D. N. (2022). GLOBathy, the global lakes bathymetric dataset. *Scientific Data*, *9*(36), 36. https://doi.org/10.1038/s41597-022-01132-9

Kilpatrick, K. A., Podestá, G., Walsh, S., Williams, E., Halliwell, V., Szczodrak, M., et al. (2015). A decade of sea surface temperature from MODIS. *Remote Sensing of Environment*, *165*, 27–41. https://doi.org/10.1016/j.rse.2015.04.023

Knauss, J. A., & Garfield, N. (2016). *Introduction to physical oceanography*. Waveland Press.

Kourzeneva, E., Asensio, H., Martin, E., & Faroux, S. (2012). Global gridded dataset of lake coverage and lake depth for use in numerical weather prediction and climate modelling. *Tellus A: Dynamic Meteorology and Oceanography*, *64*(1), 15640. https://doi.org/10.3402/tellusa.v64i0.15640

Lawrence, D. M., Fisher, R. A., Koven, C. D., Olesen, K. W., Swenson, S. C., Bonan, G., et al. (2019). The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018ms001583

Lenters, J., Anderton, J., Blanken, P., Spence, C., & Suyker, A. (2013). *Assessing the impacts of climate variability and change on great lakes evaporation* 2011 Project Reports. Great Lakes Integrated Sciences and Assessments (GLISA) Center.

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, *99*(D7), 14415–14428. https://doi.org/10.1029/94jd00483

Madonna, F., Essa, Y. H., Marra, F., Serva, F., Gardiner, T., Sarakhs, F. K., et al. (2023). Uncertainties on climate extreme indices estimated from us climate reference network (USCRN) near-surface temperatures. *Journal of Geophysical Research: Atmospheres*, *128*(11), e2022JD038057. https://doi.org/10.1029/2022jd038057

Martynov, A., Sushama, L., & Laprise, R. (2010). Simulation of temperate freezing lakes by one-dimensional lake models: Performance assessment for interactive coupling with regional climate models. *Boreal Environment Research*, *15*(2), 143.

Militino, A., Moradi, M., & Ugarte, M. (2020). On the performances of trend and change-point detection methods for remote sensing data. *Remote Sensing*, *12*(6), 1008. https://doi.org/10.3390/rs12061008

Minallah, S., & Steiner, A. L. (2021). The effects of lake representation on the regional hydroclimate in the ECMWF reanalysis. *Monthly Weather Review*, *149*(6), 1747–1766.

Notaro, M., Bennington, V., & Lofgren, B. M. (2015). Dynamical downscaling-based projections of Great Lakes water levels. *Journal of Climate*, *28*(24), 9721–9745. https://doi.org/10.1175/jcli-d-14-00847.1

Novick, K. A., Ficklin, D. L., Stoy, P. C., Williams, C. A., Bohrer, G., Oishi, C. A., et al. (2016). The increasing importance of atmospheric demand for ecosystem water and carbon fluxes. *Nature Climate Change*, *6*(11), 1023–1027. https://doi.org/10.1038/nclimate3114

O'Reilly, C. M., Sharma, S., Gray, D. K., Hampton, S. E., Read, J. S., Rowley, R. J., et al. (2015). Rapid and highly variable warming of lake surface waters around the globe. *Geophysical Research Letters*, *42*(24), 10–773. https://doi.org/10.1002/2015gl066235

Raumann, C. G., Stine, S., Evans, A., & Wilson, J. (2002). In *Digital bathymetric model of Mono Lake, California (Version 1.0)* (Tech. Rep.). U.S. Geological Survey, Coastal Marine Geology. https://doi.org/10.3133/mf2393

Sharma, S., Filazzola, A., Nguyen, T., Imrit, M. A., Blagrave, K., Bouffard, D., et al. (2022). Long-term ice phenology records spanning up to 578 years for 78 lakes around the northern hemisphere. *Scientific Data*, *9*(1), 318. https://doi.org/10.1038/s41597-022-01391-6

Smirnova, T. G., Brown, J. J., Benjamin, S. G., & Kenyon, J. P. (2016). Modifications to the rapid update cycle land surface model (RUC LSM) available in the weather research and forecasting (WRF) model. *Monthly Weather Review*, *144*(8), 1851–1865. https://doi.org/10.1175/mwr-d-15-0198.1

Sorensen, T., Espey, E., Kelley, J. G. W., Kessler, J., & Gronewold, A. D. (2024). A database of in situ water temperatures for large inland lakes across the coterminous United States. *Scientific Data*, *11*(1), 282. https://doi.org/10.1038/s41597-024-03103-8

Spence, C., Blanken, P. D., Lenters, J. D., & Hedstrom, N. (2013). The importance of spring and autumn atmospheric conditions for the evaporation regime of Lake Superior. *Journal of Hydrometeorology*, *14*(5), 1647–1658. https://doi.org/10.1175/jhm-d-12-0170.1

Subin, Z. M., Riley, W. J., & Mironov, D. (2012). An improved lake model for climate simulations: Model structure, evaluation, and sensitivity analyses in CESM1. *Journal of Advances in Modeling Earth Systems*, *4*(1). https://doi.org/10.1029/2011ms000072

Wright, D. M., Posselt, D. J., & Steiner, A. L. (2013). Sensitivity of lake-effect snowfall to lake ice cover and temperature in the Great Lakes region. *Monthly Weather Review*, *141*(2), 670–689. https://doi.org/10.1175/mwr-d-12-00038.1

Xiang, T., Vivoni, E. R., Gochis, D. J., & Mascaro, G. (2017). On the diurnal cycle of surface energy fluxes in the North American monsoon region using the WRF-Hydro modeling system. *Journal of Geophysical Research: Atmospheres*, *122*(17), 9024–9049. https://doi.org/10.1002/2017jd026472

Xiao, C., Lofgren, B. M., Wang, J., & Chu, P. Y. (2016). Improving the lake scheme within a coupled WRF-lake model in the Laurentian Great Lakes. *Journal of Advances in Modeling Earth Systems*, *8*(4), 1969–1985. https://doi.org/10.1002/2016ms000717

Xiao, K., Griffis, T. J., Baker, J. M., Bolstad, P. V., Erickson, M. D., Lee, X., et al. (2018). Evaporation from a temperate closed-basin lake and its impact on present, past, and future water level. *Journal of Hydrology*, *561*, 59–75. https://doi.org/10.1016/j.jhydrol.2018.03.059

Yao, F., Livneh, B., Rajagopalan, B., Wang, J., Crétaux, J.-F., Wada, Y., & Berge-Nguyen, M. (2023). Satellites reveal widespread decline in global water storage. *Science*, *380*(6646), 743–749. https://doi.org/10.1126/science.abo2812

Zhan, P., Song, C., Liu, K., Chen, T., Ke, L., Luo, S., & Fan, C. (2023). Can we estimate the lake mean depth and volume from the deepest record and auxiliary geospatial parameters? *Journal of Hydrology*, *617*, 128958. https://doi.org/10.1016/j.jhydrol.2022.128958

Zhang, Q., Jin, J., Wang, X., Budy, P., Barrett, N., & Null, S. E. (2019). Improving lake mixing process simulations in the Community Land Model by using K profile parameterization. *Hydrology and Earth System Sciences*, *23*(12), 4969–4982. https://doi.org/10.5194/hess-23-4969-2019