
Refining Long-Term Analysis of AVHRR Surface Temperatures for the Laurentian Great Lakes

Charles H. White

A thesis submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE

(Atmospheric and Oceanic Sciences)

at the

UNIVERSITY OF WISCONSIN-MADISON

2018

Abstract

Refining Long-Term Analysis of AVHRR Surface Temperatures for the Laurentian Great Lakes

by Charles H. White

Inland waters across the globe have been observed to be warming more quickly than their regional air temperatures, and at variable rates. Variable warming is partially attributable to individual lake geography and morphology. In very large lakes, there is also significant intra-lake variability both in surface temperature and its long-term trends. Thus, neither lake-wide mean temperatures, nor a single mid-lake sampling station are fully representative of the entire lake surface. While long-term *in situ* records are essential as benchmarks for analyzing water temperatures, these records are relatively scarce, and are typically available only from a single location per lake. Satellite data can supplement *in situ* records by offering spatially resolved views of the lake surface, albeit with gaps due to cloud interference. To date, analyses of lake surface temperature from satellite have largely focused on long-term trends for a specific portion of the year, and for a single location on the lake surface.

In light of the heterogeneity of lake surface water temperature (LSWT) and LSWT trends in large lakes, we revisit the Advanced Very High Resolution Radiometer (AVHRR) record for the Laurentian Great Lakes using a long-term (1986-2016) and high resolution (0.018°) dataset to diagnose spatial and temporal variability. While satellite data has the significant benefit of extensive spatial coverage, there are several hindrances that we aim to overcome in developing our lake surface temperature time series. Specific issues include

accounting for diurnal variability in LSWT, georegistration errors, and gap filling under cloudy skies.

After addressing these issues we give an updated perspective on the long-term mean trends in LSWT in this region for the ice-free season. Additionally we attempt to quantify the spatiotemporal variability in lake surface temperature by finding Empirical Orthogonal Functions (EOFs) to the time series. An upwelling identification method is also presented to document the duration and frequency of large upwelling events in the Laurentian Great Lakes. These idiosyncratic events are responsible for a large portion of the spatiotemporal variability in lake surface temperatures. By accounting for both long-term trends and episodic events, our analyses offer insight into the shifting thermal dynamics of these irreplaceable inland seas.

Contents

List of Figures	iv
List of Tables	viii
1. Introduction	1
2. Background	6
3. Datasets	10
4. Methodology	
4.1 Geolocation Corrections.....	12
4.2 Cloud, Ice, and Land Masking.....	13
4.3 Sea Surface Temperature Calculation.....	14
4.4 Diurnal Correction and Compositing.....	15
4.5 Gap Filling.....	16
5. Results	
5.1 Geolocation Corrections.....	19
5.2 Cloud, Ice and Land Masking.....	21
5.3 Sea Surface Temperature Calculation.....	22
5.4 Diurnal Correction and Compositing.....	26
5.5 Gap Filling Cross-Validation.....	31
6. Analysis of the Record	
6.1 Mean Intra-annual Variability.....	34
6.2 Long-Term Trends.....	39
6.3 Spatiotemporal Variability.....	43
6.4 Upwelling.....	45
7. Discussion	52
8. Conclusions	68
9. Bibliography	70
10. Appendix	76

List of Figures

- Figure 1.** AVHRR observation times of each AVHRR sensor over Lake Ontario. Each line represents the 100-day rolling mean in scan-line time for the specified sensor..... 10
- Figure 2.** Example of the original and adjusted georegistration of a channel 4 brightness temperature (K) HRPT image over Lake Michigan. The red outline indicates the actual coastline. A zonal offset of -4 grid cells, and a meridional offset of -5 grid cells is made between the original and the corrected image. 19
- Figure 3.** Frequency of georegistration corrections applied to each of the regrided HRPT images in this record expressed in units of 0.018° grid cells 20
- Figure 4.** Percent of surface of lake Ontario obscured by clouds (blue), covered by ice (orange), and free of both ice and clouds (green). The values indicated by the lines represent the average fraction of all overpasses in each month averaged for the whole record. Values do not necessarily sum to 100% due to viewing angles $>45^\circ$ being discarded for the mean cloud-free and ice-free amount, that aren't discarded when computing mean cloud or ice cover, and the coincidence of ice-cover and cloud-cover. 22
- Figure 5.** Comparison of the AVHRR LSWT measurements to their corresponding in situ buoy matchups for all platforms combined. Colored is the point density of observations with cooler colors (purple) indicating the lowest point density and warmer colors (red) indicating the highest point density. The dashed line indicates 1 to 1 correspondence. 23
- Figure 6.** Yearly averaged RMSE for all sensors (black), and each individual sensor (colored) plotted for the lifetime of each sensor..... 24
- Figure 7.** Differences between the AVHRR measured LSWT and the buoy water temperature. Filled circles indicate the mean bias, the error bars indicate the standard deviation of the errors, and the overlaid histogram represents the percent of observations occurring at each wind speed in 1 m/s bins centered on each value. Wind speeds are measured at 3-meters on the buoys. 25
- Figure 8.** Broad agreement between SST estimates based on the NLSST and MCSST equations applied to AVHRR data from 1986-2016. Bin widths are 0.0025 K. Total $n=18,680$ 26
- Figure 9.** Average diurnal cycles found in the buoy observations for the Laurentian Great Lakes plotted as a function of departure from the daily average temperature. The blue lines indicate the mean for individual years for the specified month, while the black line indicates the mean across all years. The range indicates the difference between the maximum and minimum values of the mean across all years. 27

- Figure 10.** Average diurnal cycles found in the AVHRR observations for April through November. The blue points indicate the individual differences calculated between an observation at the specified time and the observation in the reference window (8-11 UTC). The black line indicates the diurnal cycle function fitted to the centered differences. The range indicates the difference between the maximum and minimum values of diurnal cycle function.....28
- Figure 11.** Mean differences between AVHRR SST and buoy water temperature in June, July and August. Each filled circle represents the mean error for the 3 hour window centered on the circle.....29
- Figure 12.** (a) Number of all-sky views of Lake Ontario, and (b) the number of calculated daily composites.....30
- Figure 13.** The calculated mean squared errors between a location (a) near the western shoreline of Lake Michigan, and (b) towards the middle of Lake Michigan plotted for time lags of -6,-3,0,3, and 6 days in July. The grey point indicates the location from which mean squared errors from all other locations are calculated to.....31
- Figure 14.** Same as Figure 13, but for the month of February for a mid-lake point. For some locations, values are not computed due to an infrequency of clear-sky and ice-free observations of the lake during the winter.....32
- Figure 15.** Climatology of the mean lake temperature, and the range of temperatures in each of the five lakes. The thick line represents the 31-year climatology of lake-wide mean temperature. The top and bottom of the shaded areas represent the climatology of the 95th and 5th percentile of the lake-surface temperature distribution across each lake. Julian days between 0 and 50 are not shown due to frequent ice cover.....34
- Figure 16.** Climatological mean rate of change of lake-wide mean surface temperature (1986-2016). The rate of change is calculated by a 30-day moving linear regression.....35
- Figure 17.** Climatological (1986-2016) mean intra-annual rate of change of the shallowest 10% (dashed line), and the deepest 10% (solid line) of each lake. The rates of change are calculated by a 30-day moving linear regression.....36
- Figure 18.** Climatological standard deviation of temperature across the lake surface for a given day of the year.37
- Figure 19.** Interannual variation of mean lake temperature. The values of indicated by each line indicate the standard deviation of 31 mean lake temperatures (1 per year) for each day of the year for five lakes.37
- Figure 20.** Bathymetric and thermal properties of the Laurentian Great Lakes. (a) Bathymetry, with increasing positive values indicating deeper locations (b) Mean Julian day

of crossing the 4°C threshold (c) Mean Julian day of reaching 31-year mean maximum temperature.	38
Figure 21. 31-year (1986-2016) linear regression slopes for mean July-August-September (JAS) surface temperatures plotted in degrees Celsius per decade.....	39
Figure 22. Trends in monthly mean temperature for Lake Michigan. Trends are only shown for months and locations where more than 28 years are ice-free.	40
Figure 23. Surface temperature trends (1986-2016) across all Laurentian Great Lakes for each month of the year. The orange line indicates the median trend, box edges indicate the 25 th and 75 th percentile, and the whiskers indicate the last values within one interquartile range of the 25 th and 75 th percentiles.	41
Figure 24. Same as Figure 23, but each box is shown for individual lakes.	41
Figure 25. Pearson correlation coefficient between the monthly 1986-2016 trends, and the (a) climatological mean monthly temperature and (b) bathymetry.....	42
Figure 26. Trends in stratification date as a function of lake depth. Values indicated by the black and colored lines represent means calculated in 7.14 meter intervals. A negative trend implies stratification is occurring earlier in the year with respect to the mean date.....	43
Figure 27. The first four empirical orthogonal functions (EOFs) calculated for the deseasonalized and detrended surface temperature time series. The EOFs are plotted as a function of correlation coefficient between their corresponding principle component, and the surface temperature time series at each grid point. Each EOF is calculated independently for each lake.	44
Figure 28. Surface temperature image from Jun 21 2015. Shown are the two points, one inshore, and one offshore, that are referenced in Figure 30 and Figure 31.....	46
Figure 29. Surface temperature time series and the upwelling tests for the inshore point indicated in Figure 29. The dashed grey lines indicate the thresholds defined in Table 5. The grey shading in (a) indicates where upwelling is occurring defined by this algorithm. The same shading in (b), (c), and (d) are where the individual tests are passed.	47
Figure 30. Same as Figure 30 but for the offshore point indicated in Figure 29.....	48
Figure 31. The number of discrete upwelling events identified. Only locations with more than five events are shown.....	50
Figure 32. The average duration, in days, of the discrete events shown in Figure 32. Only locations that had more than five documented upwelling events, and are within 5 grid cells of a shoreline are shown.	50

Figure 33. Mean wind direction in each shaded region when an increase in upwelling area occurs in a few selected regions. The wind roses indicate the direction where the wind is coming from.....	51
Figure 34. Climatological (1986-2016) mean temperature difference between the shallowest 10% and deepest 10% of each lake	56
Figure 35. JAS mean trend calculated for three periods (a) the 1986-2016 AVHRR dataset used in this work (b) the 1994-2013 window used in Mason et al. 2016 and (c) the 1992-2010 window used in Woolway et al. 2018.	58
Figure 36. Monthly mean trends (1986-2016) in lake surface temperature shown for each lake for the shallowest 10% and the deepest 10% of the five lakes.....	60
Figure 37. July-September (JAS) and September-November (SON) mean trends in lake surface temperature across the entire Laurentian Great Lakes region calculated for (a) varying start years, (b) varying end years, and (c) a rolling 20-year time series shown at the start year.	62
Figure 38. Long-term trends for Lake Michigan using the five different methods detailed in Table 4. Box edges indicate quartiles, the middle line indicates the median, and the whiskers indicate the last values within one interquartile range of the 25 th and 75 th percentiles.....	64
Figure A.1. Trends in monthly mean temperature for Lake Ontario. Trends are only shown for months and locations where more than 28 years are ice-free.....	76
Figure A.2. Trends in monthly mean temperature for Lake Huron. Trends are only shown for months and locations where more than 28 years are ice-free.....	76
Figure A.3. Trends in monthly mean temperature for Lake Superior. Trends are only shown for months and locations where more than 28 years are ice free.....	77
Figure A.4. Trends in monthly mean temperature for Lake Erie. Trends are only shown for months and locations where more than 28 years are ice-free.....	77

List of Tables

Table 1. Zonal and Meridional georegistration adjustments for each AVHRR platform. The mean and standard deviations are expressed in units of 0.018° grid cells. Values for ascending and descending nodes are expressed separately with descending nodes in parentheses.....	20
Table 2. Root mean squared error between the AVHRR LSWT, and the <i>in situ</i> water temperature reported for the entire lifetime of the sensor.....	23
Table 3. Results from the cross-validation. Root mean squared error (RMSE), mean bias, and Pearson correlation coefficient (ρ) are calculated between the three different methods used: a locally weighted average (LWA), LOESS, and a combination of selective LWA and LOESS. n=33,593.....	33
Table 4. Percent of variance in the surface temperature time series that is explained by each of the EOF modes shown with the associated errors in the eigenvalues expressed as a percent.	44
Table 5. The tests used to determine areas effected by large upwelling events, their physical reasoning for inclusion in this algorithm, and their quantitative implementation.	46
Table 6. Pearson correlation coefficient of 1986-2016 trends to the trends calculated for 1995-2013 and 1992-2010 periods using the same AVHRR HRPT dataset.....	57
Table 7. Details of the five sets of data from which long term trends are calculated in Figure 31.....	63

1 Introduction

Climate change can affect lakes both directly and indirectly, and is considered one of the largest threats to freshwater ecosystems (Carpenter et al. 2011). Due to their sensitivity to climate, in some respects, lakes can serve as sentinels of climate change (Adrian et al. 2009; Schindler 2009). Globally, lake surface water temperatures (LSWTs) are warming at highly variable rates (O'Reilly et al. 2015), with high latitude lakes generally warming the fastest (Schneider and Hook 2010). The warming water temperatures can have profound effects on freshwater ecosystems such as declining oxygen concentrations (Blumberg and Di Toro 1990), increasing the basal metabolic rates of freshwater species (Gillooly et al. 2001; Kraemer et al. 2017), and facilitating the spread of invasive species (Rahel and Olden 2008; Smith et al. 2012). The physical impacts experienced by inland waters as a result of lake warming include changes in wind speeds over lakes (Desai et al. 2009), decreases in ice cover (Magnuson et al. 2000; Austin and Colman 2007), and changes in stratification of the water column (Kraemer et al. 2015; Winslow et al. 2017). Given the previously observed physical changes, and ecological implications in global lakes, a long-term monitoring system is essential to diagnosing the directionality, magnitude, and variability of water temperature trends. While long term *in situ* records of water temperature are ideal for detecting such changes, these records are not available for most of the world's lakes and are usually lacking significantly in spatial coverage of any one lake.

Long-term satellite records of infrared imagery have the ability to overcome this data limitation, as illustrated in several recent global-scale studies (O'Reilly et al. 2015; Schneider and Hook 2010; Mason et al. 2016; Woolway and Merchant 2018; Pareeth et al. 2016; Riffler

et al. 2015). Schneider and Hook (2010) demonstrated rapid warming in surface waters of 167 inland water bodies, and found good agreement between satellite and *in situ* records from a few lakes. O'Reilly et al. (2015) used a mixture of *in situ* data and satellite time series to reveal heterogeneity in warming rates among lakes across the globe. O'Reilly et al. (2015) further concluded that these lakes often warm more quickly or slowly relative to air temperature, reflecting an individual lake's combination of both climate and geomorphic factors. This motivates continued regional and global scale studies of inland water temperatures since few assumptions can be made about how surface water temperature is changing with respect to air temperature trends, neighboring lakes, and other geomorphic factors.

Satellite sensors have long been used to measure sea surface temperatures (SSTs), and for time series analysis of inland waters, but long-term spatiotemporal analyses of variation within individual inland water bodies remains scarce. Quite often, such a detailed analysis is hampered due to the infrequency of clear-sky observations of the surface. Further complicating, is the larger variability associated with daytime measurements of surface temperature compared to nighttime measurements (Wilson et al. 2013). Additionally, the skin-effect and mean diurnal cycle of water temperature (Woolway et al. 2016; Wilson et al. 2013) can complicate trend analysis if there is substantial orbital drift of the sensor platform during the study period. It is suggested that nighttime observations be used to find lake surface temperature due to the increased variability during the day (Wilson et al. 2013). Because of the inconsistent availability of clear-sky, nighttime observations during the summer warm season, it is common to use averages across July, August, September (JAS) to assess lake temperature trends (Schneider and Hook 2010; O'Reilly et al. 2015; Mason et al. 2016).

In the northern hemisphere JAS represents the peak of the annual cycle of lake surface temperature, and while JAS represents a critical portion of the annual cycle, it does not tell the whole story of how climate change could affect lakes (Winslow et al. 2017). There has been large demonstrated variability in long term mean trends between individual months in JAS for individual lakes. This variability has been demonstrated to surpass the heterogeneity in warming rates between different lakes. Depending on latitude, size, bathymetry and atmospheric conditions, JAS could represent a different portion of the mean annual cycle of lake surface temperature with respect to the warming and cooling phases of different lakes in the same hemisphere.

Additionally, many studies have independently identified significant intralake spatial heterogeneity in JAS warming trends (Woolway and Merchant 2018; Mason et al. 2016; Pareeth et al. 2016). This gives motivation to attempt further investigation of lake surface temperature heterogeneity outside of the typical JAS seasonal average at sub-seasonal temporal scales and relatively fine spatial scales. In very large lakes, spatial variability in warming rates have been suggested to arise from a number of factors. Woolway and Merchant (2018) investigated the relationship between intra-lake variability in long-term trends with bathymetry and the date at which thermal stratification occurs. They identified correspondence between warming rates and bathymetry with the deeper areas of the many lakes warming more quickly than shallower areas in large lakes in the Northern Hemisphere. Others have attributed similar changes in the Laurentian Great Lakes to milder winter conditions, and a combination of meteorological factors resulting in earlier springtime stratification particularly in the deepest areas (Zhong et al. 2016).

In the Laurentian Great Lakes region, several *in situ* studies have recognized the significance of surface temperature heterogeneity particularly due to coastal upwelling. Haffner et al. (1984) observed upwelling events and investigated their ability to transport nutrients towards the surface in Lake Ontario, which in turn led to rapid shifts in the composition of zooplankton communities. Coupled general circulation and ecosystem models have identified upwelling as a major mechanism to supply hypolimnetic nutrients and generate significant productivity after offshore thermal stratification (Pilcher et al. 2015). Rao and Murthy (2001) made measurements of upwelling in Lake Ontario documenting upwelling episodes that lasted approximately 4-6 days. These events occurred under the influence of prevailing winds, with additional evidence for the influence of internal Kelvin waves on upwelling. Using satellite data from the Great Lakes Surface Environmental Analysis (GLSEA; Schwab et al. 1999), examples of large upwelling events have been characterized for Lake Michigan in the 1992 to 2000 period (Plattner et al. 2006).

The goal of this work is to leverage the length (31-years) and spatial resolution of the 1.1 km Advanced Very High Resolution Radiometer (AVHRR) High Resolution Picture Transmission (HRPT) and Full Resolution Area Coverage (FRAC) record to provide a refined long-term analysis of surface temperatures in the Laurentian Great Lakes. This dataset will provide a spatially resolved daily view of the five largest lakes in the region. In this analysis, we will describe the climatology of LSWT in the Laurentian Great Lakes. Intralake variability in surface temperature will be quantified on a sub-seasonal basis with respect to the mean annual cycle, and interannual changes. Given the ecological significance of surface temperature variation, we will identify the main modes of variability in lake surface temperature. In particular, we will attempt to identify and document the location, size and

frequency of large upwelling events that are large contributors to spatiotemporal variation of surface temperature for the largest five lakes in the region. By accounting for both long-term trends and episodic events, our analyses offer insight into the shifting thermal dynamics of these irreplaceable inland seas.

2 Background

There are several satellite sensors capable of measuring sea surface temperatures accurately at high resolution, but there are a select few with records long enough for climatological analysis. The longest of these for the Laurentian Great Lakes region begins with the AVHRR HRPT data available continuously from NOAA-9 starting in 1986. AVHRR is a space-borne visible and infrared imager on the NOAA Polar-orbiting Operational Environmental Satellite (POES) project platforms and the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Meteorological Operational (MetOp) satellites. The native spatial resolution of the sensor is 1.1km at nadir, with a total swath width of 2900km. All AVHRR-carrying satellites are in sun-synchronous orbits providing at least twice-daily observations for nearly all locations on earth

The HRPT data are available only for the locations around receiving stations at the native 1.1 km resolution of the sensor. The HRPT record for the North American Great Lakes is available on a continuous basis starting in 1986 with NOAA-9, and only sparsely on previous platforms for this region. On MetOp satellites, the FRAC data is of 1.1 km resolution, like the HRPT, but is available globally starting in late 2006 with the launch of MetOp-A. Altogether, the AVHRR record is the longest satellite record using a consistent set of imagers with similar channel selection and bandwidth. A continuous 1.1 km record of the Laurentian Great Lakes region can be assembled by combining the regionally available HRPT, with a regional cut-out of the globally available FRAC.

Since the beginning of the satellite era there has been active interest in the derivation and validation of sea surface temperature (SST) measurements, particularly from AVHRR. In

the infrared regime, there exists an atmospheric window, where continuum water vapor absorption is very low. In the early 1970s, the use of two infrared wavelengths in the infrared window region to correct for the effects of water vapor absorption in this window was investigated (Anding and Kauth 1970; Maul and Sidran 1972; Prabhakara et al. 1974). It was found that SST could be accurately determined in the cloud free atmosphere using a pair of wavelengths that exhibited differential absorption due to the presence of water vapor.

After some refinement for moist atmospheres (McMillin 1975), several studies demonstrated the accuracy of the split window approach for the AVHRR on NOAA-7 and NOAA-9 to be within 1 K (McClain et al. 1983, 1985; McMillin and Crosby 1984). A history of the development of this approach is reviewed in (McMillin and Crosby 1984). The result, is the multiple-channel sea surface temperature (MCSST) stated simply as

$$MCSST = T_j + \gamma(T_i - T_j) \quad [1]$$

where T_i and T_j are the brightness temperatures measured at two wavelengths. The gamma parameter accounts for the differential water vapor absorption between the two channels. The $\gamma(T_i - T_j)$ term acts as a water vapor correction, where gamma is assumed to be constant. To correct for temperature-dependent bias and to better capture the bulk SST rather than the skin, a regression form of the MCSST equation is used as shown in equation 2.

$$MCSST = b_1T_i + b_2(T_i - T_j) - b_3 \quad [2]$$

The coefficients can be determined by a regression to *in situ* measurements by ships, moored buoys or drifting buoys (Strong and McClain 1984).

Further investigating the accuracy of the MCSST, Walton (1988) noted a temperature and water vapor dependence of γ , and presented a nonlinear multichannel algorithm. This

nonlinear algorithm was later simplified and formalized as the Nonlinear Sea Surface Temperature (NLSST) as in equation 3 (Walton et al. 1998) where the gamma parameter γ is multiplied by an estimate of the surface temperature T .

$$NLSST = T_i + \gamma(T)(T_i - T_j) \quad [3]$$

A correction for satellite viewing angle θ (Llewellyn-Jones et al. 1984) can be added to the MCSST and the NLSST to produce equations 4 and 5, for which the coefficients can be fitted to *in situ* measurements.

$$MCSST = b_1 T_{11\mu m} + b_2 (T_{11\mu m} - T_{12\mu m}) + b_3 (T_{11\mu m} - T_{12\mu m})(\sec \theta - 1) + b_4 \quad [4]$$

$$NLSST = a_1 T_{11\mu m} + a_2 (T_{11\mu m} - T_{12\mu m})(SST_{guess}) + a_3 (T_{11\mu m} - T_{12\mu m})(\sec \theta - 1) + a_4 \quad [5]$$

SST_{guess} is an initial estimate of the sea surface temperature used to adjust the gamma parameter in equation 3. It is suggested in Walton et al. (1998) that this initial estimate could come from climatology or the MCSST. Kilpatrick et al. (2001) noted that although these coefficients can be fit globally, it may improve accuracy to fit the equations for two different water vapor regimes. They also noted that accuracy improved by fitting coefficients within a time-dependent window. SST measurements have also been validated for lakes and coastal regions under a variety of conditions documenting error standard deviations of about 1 K, and that morning fog in the lakes contributed significantly to the measurement error (Li et al 2001).

When measuring SST, the depth at which the measurement is made is critical information due to the amount of variability that can occur in shallow temperature profiles (Donlon et al 2002). Particularly, distinction is made between the skin temperature measured by infrared radiometers, and bulk temperatures often made by *in situ* measurements. The difference between these two sets of measurements is referred to as the skin-effect (Schluessel et al. 1990; Minnett et al. 2011), and can have significant impact on interpretation of the

measurement due to thermal stratification in the water column. Oceanographers are often more interested in bulk measurements of surface temperature, hence the regression of SST coefficients to *in situ* measurements of bulk temperature. However, the differences between the skin and the bulk temperatures can add uncertainty to the measurements (Donlon et al. 2002). In the ocean, these diurnal differences are largely a function of solar insolation and are lessened by mixing as wind speeds increase (Webster et al. 1996; Gentemann et al. 2003; Minnett et al. 2011). At low wind speeds, diurnal warming can elevate daytime surface temperature by 5 to 7 K above nighttime observations (Gentemann et al. 2008).

While most work on identifying diurnal variability in SSTs has been performed using satellite and *in situ* measurements over the ocean, much of what has been identified is also applicable to lakes. Wilson et al. (2013) noted that cloud-cover, high humidity, and wind speed can impact the magnitude of the lake skin-effect and that the lake skin-effect was more variable in daytime observations. While the oceanic skin-effect and the lake skin-effect are mechanistically the same, lakes undergo a wider range of atmospheric conditions that can drive skin-bulk differences (Wilson et al. 2013). Woolway et al. (2016) identified an association between lake size and summer diurnal temperature range which could amount to 7 K for small lakes, and less than 2 K for very large lakes. For polar orbiting satellites, it is plausible that orbital drift could introduce bias to long-term analyses of lake surface temperature by failing to control for diurnal temperature swings. This could be particularly important for calculating long term trends using records from multiple satellites that view lakes at different periods of the day. Additionally, day-to-day variability could be affected in cases where measurements from successive days reflect a significantly different portion of the diurnal cycle.

3 Datasets

The Pathfinder Atmospheres-Extended (PATMOS-x) dataset (Heidinger et al. 2014) is a global climate data record made up of the over 36 year-long AVHRR record (1982 to present). In PATMOS-x, the AVHRR GAC data are subsampled to a regular global 0.1° latitude/longitude grid. For each day, the ascending and descending nodes are composited to typically provide one daytime, and one nighttime global view per sensor. For this study, a high resolution version of the PATMOS-x data set was produced on a regular 0.018° latitude/longitude grid (roughly 2 km x 1.5 km in this region) from the HRPT and FRAC datasets available over the North American Great Lakes region from 1986 through 2016. This resolution was chosen due to the decreasing pixel density at larger viewing angles. The high resolution dataset was produced for the region bounded by the longitudes 95°W to 73°W , and the latitudes 40°N to 50°N . The specific satellite platforms used in this record all experience significant orbital drift over their lifetime, with the exception of MetOp-1, and MetOp-2 (Figure 1).

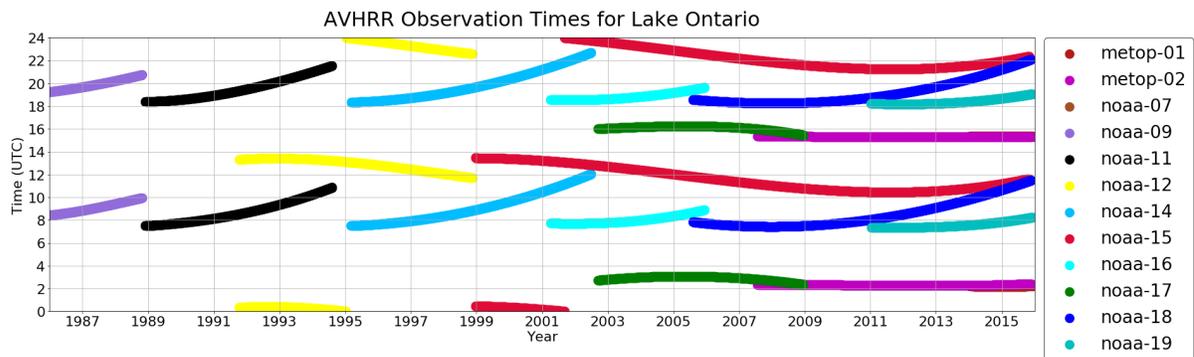


Figure 1. AVHRR observation times of each AVHRR sensor over Lake Ontario. Each line represents the 100-day rolling mean in scan-line time for the specified sensor.

The Great Lakes Ice Cover Database is a dataset of ice concentration over the Laurentian Great Lakes since 1973. These observed and interpolated data come from a blend

of sources. The data from 1973-2005 (Assel 2003, 2005) are provided on an approximate 2.5 km grid. The 2006 to 2013 data are provided on a 1.275 km grid (Wang et al. 2012). For this study, both of these datasets were nearest-neighbor interpolated in space and time to the 0.018 x 0.018 degree grid at a daily time step. These data are used to remove surface temperatures that may be contaminated with significant ice cover where LSWT measurements may have low confidence.

In situ measurements of water temperature and other buoy measurements are taken from the *In situ* SST Quality Monitor (iQuam) (Xu and Ignatov 2014). The *iQuam* dataset includes observations from several types of platforms including ships, drifters, and moored buoys. *iQuam* implements additional quality control measures to ensure the accuracy of the *in situ* measurements from these platforms. This is essential for satellite to buoy matchups, since a few poor measurements can significantly affect the fitted SST coefficients. The *in situ* data from the iQuam dataset provides hourly measurements of water temperature measured at 60 cm or 100 cm depth from 8 buoys in the Great Lakes that span the entire 31-year record.

The North American Regional Reanalysis (NARR) is a high-resolution land surface hydrology dataset for North America (Mesinger et al. 2006). In this work, we use the daily 10-meter wind product provided in the NARR. These winds are interpolated from the 32 km NARR grid to our regular 0.018° grid.

4 Methodology

4.1 Geolocation Corrections

After the initial processing of the level-1b AVHRR files obtained from NOAA CLASS, there remains a number of adjustments to the dataset that need to be made. There are well documented errors in the navigation of the AVHRR sensors, creating some uncertainty in the georegistration of the AVHRR images. Bordes et al. (1992) assessed these errors by identifying coastal landmarks over Europe and documented mean errors of 3 pixels in the cross-track direction, and -0.8 scanlines in the along-track direction. There are two basic sources of navigation errors (Baldwin and Emery 1993). The first, is associated with the oblateness of the earth and the geometric distribution of measurements from AVHRR. The second source of error comes from uncertainty in the relative positioning of the satellite. Some proposed methods of correcting these errors involve using better defined orbital parameters and ground control points (GCPs). Using these methods, some have been able to improve the AVHRR navigation to within one-pixel accuracy (Moreno and Melia 1993; Baldwin and Emery 1993). These navigation errors are most problematic along regions with well-defined boundaries, like coastlines. However, since analysis performed here involves subsampling the native resolution AVHRR to a coarser grid and coastline-adjacent pixels are removed in proceeding steps, native resolution sub-pixel accuracy is not necessarily required. Therefore, a simple first-order correction involving zonal and meridional linear displacements is applied to the time series. The method used is nearly identical to that of Schwab et al. (1999) with a significantly relaxed requirement on maximum linear displacements.

This simple correction is carried out by first calculating the Laplacian of the cloud-cleared channel 4 (10.3 μm -11.3 μm) brightness temperatures. The resulting 2-D array, hereafter the edge-enhanced image, reports high values along the well-defined coastlines. A similar calculation is performed for the land classification variable included in PATMOS-x which is based on the MOD44W dataset (Carroll et al 2017) to create a reference image. Where changes in land classification occur, the pixel value is set to one, otherwise it is set to zero. A dot product is taken between the edge-enhanced image and the reference image at various linear displacements in the zonal and meridional directions. The combination of displacements that maximizes the dot product between the two images is recorded, and used to offset the original brightness temperatures and cloud mask. This method of addressing the uncertainties in the AVHRR navigation performs best in clear-sky regions where coastlines are clearly visible. When either 95% of the land or water in a given image is classified as cloudy, no adjustment is made due to the likely small amount of coastlines observed.

4.2 Cloud, Ice, and Land Masking

Cloud masking is a critical procedure when calculating SST. Clouds negatively affect the derivation of the SST coefficients, and thus the accuracy of SST products. In this work, clouds are removed using a naïve Bayesian cloud detection scheme (Heidinger et al. 2012) that is included in the PATMOS-x dataset. The algorithm produces a pixel-level naïve Bayesian probability of cloudiness. For our purposes, all pixels with >5% cloud probability were removed. Pixels that were not classified as water bodies were removed using the land classification variable. Lake ice coverage exceeding 20% in the Great Lakes Ice Atlas is

discarded and pixels that are immediately adjacent to those exceeding 20%, and within one day of exceedance are also removed. This is done in an effort to mitigate issues introduced by interpolation to the grid used in this analysis and to give conservative estimates of the ice-free lake surface.

4.3 Lake Surface Water Temperature Calculation

LSWT is calculated using a nonlinear split-window equation using channels 4 (10.3 μm -11.3 μm) and 5 (11.5 μm -12.5 μm) from the AVHRR sensor, and the sensor viewing angle. This is found by the NLSST (equation 5), for which, the MCSST (equation 4) is used as the initial guess. For this analysis, coefficients are fitted by a least-squares regression to moored buoy measurements in the Laurentian Great Lakes. The AVHRR measurements are collocated in space (nearest pixel) and time (<15 min) to buoy measurements. A separate set of coefficients is fit for each year of the record for each AVHRR sensor. Others have found increased accuracy using smaller temporal windows for fitting coefficients (Kilpatrick et al. 2001). We fit coefficients on a yearly basis because the *in situ* measurements, made in off-shore locations, do not experience the same range in temperatures as the in-shore locations in smaller temporal windows.

4.4 Diurnal Correction and Compositing

To reduce the bias associated with the time of day a measurement was made, a diurnal correction is applied to the calculated LSWT. The goal of this correction is to remove the bias from these data associated with the mean diurnal variability. The method described in Foster and Heidinger (2013), is used in our analysis with some modification in the selection of the data used to fit the diurnal cycle.

To empirically determine the timing and amplitude of the mean diurnal cycle, differences at all times of day are calculated between all LSWT values and the LSWT at some reference time. For all days, provided that there are two clear sky observations, and one is in the reference time window, the difference is recorded. The differences are centered around zero, and are used to fit the coefficients of the diurnal cycle function of Foster and Heidinger (2013). The diurnal cycle is modeled by a two-harmonic sine function. The reference was chosen to be the LSWT measured between 8UTC and 11UTC since there is an AVHRR overpass in this window each day for nearly the entire record. The diurnal cycle in this window is also relatively flat compared to other windows during the day.

The observations are discarded if the surrounding 20x20 LSWT array is more than 50% cloudy or if the standard deviation of the 20x20 array differs from that of the reference by more than 1 K. This calculation is performed at all buoy locations for each month of the year even when the buoy was not present. This is done to ensure a variety of weather conditions were considered and that the measurements are not affected by near-shore variability. Other methods were attempted, such as only considering days where at least 6 clear-sky observations were obtained and subtracting out the average. These other methods resulted in diurnal cycles with

shapes that were largely inconsistent with buoy measurements likely due to relatively small amount of cases with entirely clear skies for the whole day. In contrast, the method used here only requires two clear-sky observations with one occurring during the reference time window to be included in the diurnal cycle fit.

4.5 Gap Filling

After the daily compositing, there still existed frequent gaps in the daily LSWT field. These gaps were typically over areas of the lakes that have experienced persistent cloudiness, but sometimes occur because the data were not available for that day from the NOAA CLASS archive. Some previous studies have aimed to estimate the surface temperatures using purely time-series based interpolation techniques like LOESS (Schneider and Hook 2010; Riffler et al. 2015), or harmonic analysis (Pareeth et al. 2016). However, there are some unique cases in this dataset where spatial information can improve the interpolated estimates. These cases involve situations where a given time series at a pixel has a significant multi-day gap, but a neighboring or nearby pixel does not have an identical gap. Most frequently, this occurs along cloud or AVHRR swath edges.

The goal of this interpolation scheme is to give greater weight to values that, in a climatological context, have small differences compared to the location of the missing value. Taking advantage of both spatial and temporal information to make estimates of LSWTs in these gaps, a weighted average of surrounding neighbors in a local spatiotemporal neighborhood is used. First, the seasonality is removed from the time series using a smoothed average across all years for each Julian day. When a missing value is encountered, a search is

made for values in a local spatiotemporal neighborhood. Here, only values within a radius of 13 grid cells, and 5 days are considered. The missing value is replaced by a weighted average of its neighbors. The average is weighted by the inverse of the climatological mean squared error at the corresponding spatial and temporal lag from the origin described in equation 6.

$$w_{x,y,t} = \frac{1}{n} \sum_0^n \frac{1}{(LSWT_o - LSWT_{x,y,t})^2} \quad [6]$$

In equation 6, $LSWT_o$ is the sea surface temperature at the location of the missing value. $SST_{x,y,t}$ is the sea surface temperature at spatial lag of x grid cells in the zonal direction, y grid cells in the meridional direction and t days before or after $LSWT_o$. n is the number of cases in the record where there are observations of both $LSWT_o$ and of $LSWT_{x,y,t}$. $w_{x,y,t}$ is the computed weight at the corresponding spatiotemporal lags. The values of w are calculated separately for each month of the year. x and y range from -13 to 13 grid cells, and t ranges from -5 to 5 days. The weights are normalized and multiplied by the neighboring deseasonalized observations the result of which, if determined to be reliable by metrics described later, replaces the missing value.

This is attempted for each occurrence of a missing value. In cases where there is not a sufficient amount of neighboring observations, or if the neighboring observations are determined to be poor estimators, the gap remained. The remaining gaps were filled using a locally weighted regression (LOESS; Cleveland 1979; Cleveland et al. 1988). A locally weighted polynomial is fit within a moving temporal window to both smooth existing observations, and predict missing ones. This local regression, along with other time series smoothing and interpolation methods has the significant disadvantage of dampening short term

variability. This dampening is somewhat reduced by performing this interpolation on a moving window of 20 days, and deseasonalizing the LSWT time series prior to interpolation.

In order to justify the computational expense of calculating the weights in equation 6 explicitly for all locations and in all months, a simple cross-validation of these methods is performed. A one-year time series of lake surface temperature observations is taken from all locations on Lake Michigan. Ten observations from each time series are removed. The removed observations are interpolated using three methods: LOESS, a locally weighted average, or a combination of both where some select values are calculated using a locally weighted average prior to using LOESS. When the combination is used, the locally weighted average estimate is only inserted in the time series where there is a large number of observations in the local neighborhood, and the standard deviation of the available observations is low. These two metrics are used to inform where we can have high confidence in the locally weighted average because they are computationally inexpensive, and indicate homogeneity in the observations (i.e. where an average might be an appropriate estimate).

5 Results

5.1 Geolocation Corrections

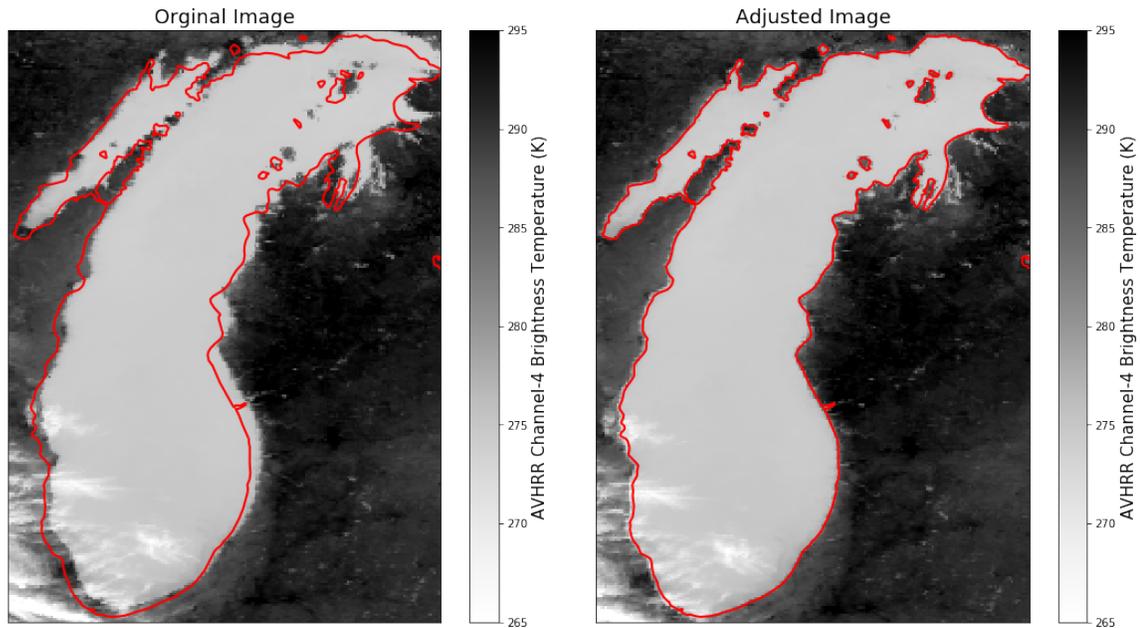


Figure 2. Example of the original and adjusted georegistration of a channel 4 brightness temperature image over Lake Michigan. The red outline indicates the actual coastline. A zonal offset of -4 grid cells, and a meridional offset of -5 grid cells is made between the original and the corrected image.

An example of an adjustment made to a single image can be seen in Figures 2a and 2b. In most cases adjustments were small, although 2.4% (3.6%) of images required a zonal (meridional) offset of the same magnitude or larger as shown in Figures 2a and 2b. The largest adjustments appear to be in the meridional direction, however the small zonal adjustments are more frequent (Figure 3). Given that the zonal distance of one grid cell (~ 1.5 km) is smaller than the meridional distance (~ 2.0 km) at this latitude, it is surprising that the meridional corrections are more often larger than the zonal ones when expressed as 0.018° grid cells. Table 1 shows the mean and standard deviation of the adjustments for each AVHRR platform. The

largest adjustments are made to the earlier sensors (NOAA-9, NOAA-11, and NOAA-12). The largest mean bias also occurs in the earlier platforms, but no platforms launched after NOAA-14 show a bias greater than a single grid cell (0.018°).

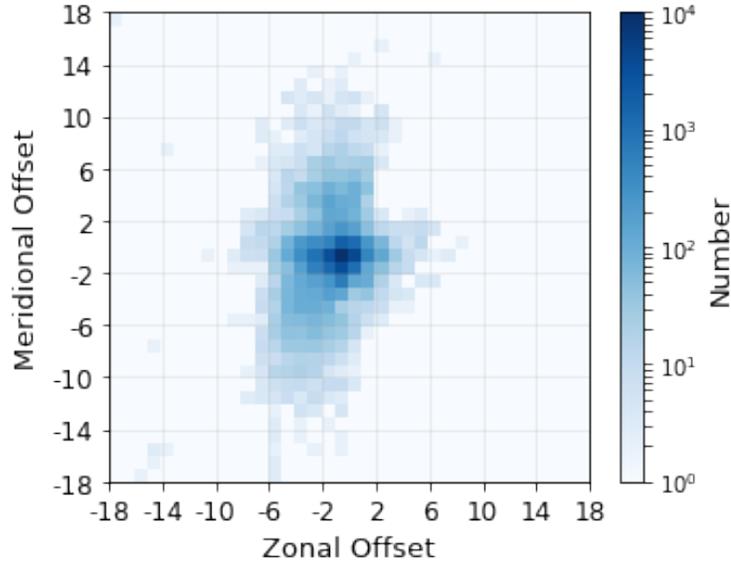


Figure 3. Number of georegistration corrections applied to each of the regrided HRPT images in this record expressed in units of 0.018° grid cells

Table 1. Zonal and Meridional georegistration adjustments for each AVHRR platform. The mean and standard deviations are expressed in units of 0.018° grid cells. Values for ascending and descending nodes are expressed separately with descending nodes in parentheses.

Platform	Number of Adjustments	Zonal Offsets		Meridional Offsets	
		Mean	Std. Dev.	Mean	Std. Dev.
NOAA-9	1205 (1117)	-0.92 (-0.80)	1.76 (1.65)	-1.51 (1.76)	2.85 (1.73)
NOAA-11	1809 (1754)	-1.23 (-0.45)	1.82 (1.49)	-2.52 (1.92)	3.15 (3.04)
NOAA-12	2355 (2694)	-1.22 (-0.19)	1.72 (1.30)	-2.37 (2.05)	3.47 (3.16)
NOAA-14	2994 (2793)	-1.21 (0.03)	1.58 (1.19)	-0.15 (0.22)	1.05 (1.10)
NOAA-15	7381 (7534)	-0.24 (0.12)	1.46 (1.52)	-0.14 (-0.14)	1.13 (1.25)
NOAA-16	2183 (1981)	-0.82 (-0.36)	1.07 (1.26)	0.06 (-0.11)	0.97 (1.10)
NOAA-17	1862 (2868)	-0.36 (-0.02)	0.91 (1.16)	-0.37 (0.28)	0.80 (0.96)
NOAA-18	4498 (4203)	-0.08 (0.18)	1.08 (1.34)	0.15 (-0.18)	1.00 (1.25)
NOAA-19	3110 (2965)	0.11 (0.19)	1.25 (1.34)	-0.19 (-0.13)	1.11 (1.28)
MetOp-A	3432 (3366)	0.18 (0.20)	1.01 (1.10)	-0.71 (0.23)	1.00 (1.17)
MetOp-B	1103 (1237)	0.27 (0.04)	1.08 (1.20)	-0.92 (0.33)	1.02 (1.17)

Mean adjustments less than a single grid cell could be attributable to the grid resampling by the PATMOS-x processing and not indicative of any navigational errors. However, in the earlier sensors, there are differences in sign of the mean zonal and meridional offsets between the ascending and descending nodes of each platform. This happens because AVHRR approaches the region from the opposite direction for each node. For platforms where the magnitude of the mean offsets of the ascending and descending nodes are similar, larger than a single grid cell, and have opposite signs, we attribute these adjustments to a mean bias in AVHRR navigation and not purely a factor of the grid resampling. This is particularly clear for the meridional offsets in NOAA-9, NOAA-11, and NOAA-12. Similarly, to some degree, the standard deviations of the adjustments can be attributed to uncertainties in our methods in cloudy scenes. However, comparison of the standard deviations between the early sensors and the more recent ones (such as NOAA-18, and NOAA-19) suggest larger variability in the georegistration of the early sensors.

5.2 Cloud, Ice, and Land Masking

As noted in (Ackerman et al. 2013), there is strong seasonality to cloudiness in the North American Great Lakes Region that has been recognized in the PATMOS-x AVHRR dataset. This leads to a larger amount of clear-sky measurements during boreal summer, and a relative scarcity of clear-sky measurements during boreal winter. As an example, Figure 4 shows the percentage of the lake surface obscured by clouds that are removed over the surface of Lake Ontario for each month averaged over all observations between 1986-2016. Together, cloud and ice masking result in a much higher observation density of LSWT during the spring and summer compared to the winter.

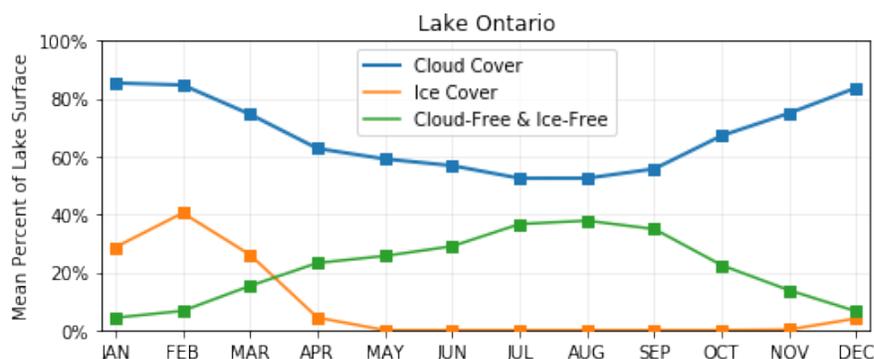


Figure 4. Percent of surface of lake Ontario obscured by clouds (blue), covered by ice (orange), and free of both ice and clouds (green). The values indicated by the lines represent the average fraction of all overpasses in each month averaged for the whole record. Values do not necessarily sum to 100% due to viewing angles $>45^\circ$ being discarded for the mean cloud-free and ice-free amount, that aren't discarded when computing mean cloud or ice cover, and the coincidence of ice-cover and cloud-cover.

5.3 LSWT Errors

The results of the LSWT coefficient fitting are shown in Table 2. The root mean squared error (RMSE) the AVHRR LSWT and the *in situ* measurements is determined to be 0.588 K across all sensors (Figure 5). Eight buoys in the Laurentian Great Lakes are used that span the entire 1986-2016 record are used in this analysis. Fewer AVHRR to buoy matchups are made between 280 K and 285 K due to relatively rapid warming and cooling that typically occurs during the spring and autumn. The RMSE of the LSWT equation changes slightly between each individual sensor. For example, earlier sensors on NOAA-12 have an RMSE as large as 0.706 K (Table 2). More recently launched sensors like NOAA-19 and MetOp-1 have a lower RMSE of 0.514 K and 0.541 K respectively. Comparing the annual mean RMSE from 1986-2016 confirms the elevated RMSE during the period covered by the first three sensors (NOAA-9, NOAA-11, and NOAA-12; Figure 6).

Table 2. Root mean squared error between the AVHRR LSWT, and the *in situ* water temperature reported for the entire lifetime of the sensor.

Platform	LSWT RMSE	Matchups
NOAA-9	0.690 K	583
NOAA-11	0.694 K	1,327
NOAA-12	0.706 K	1,387
NOAA-14	0.619 K	2,017
NOAA-15	0.586 K	6,685
NOAA-16	0.628 K	1,881
NOAA-17	0.618 K	1,523
NOAA-18	0.559 K	3,083
NOAA-19	0.514 K	3,397
MetOp-A	0.541 K	1,227
MetOp-B	0.563 K	4,547

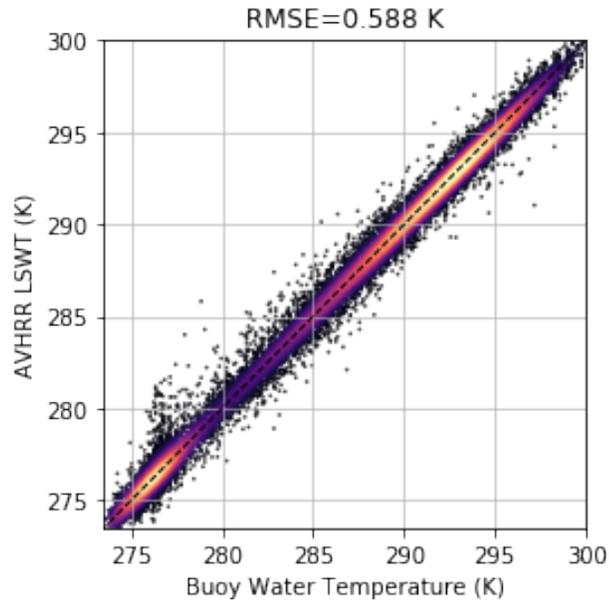


Figure 5. Comparison of the AVHRR LSWT measurements to their corresponding *in situ* buoy matchups for all platforms combined. Colored is the point density of observations with cooler colors (purple) indicating the lowest point density and warmer colors (red) indicating the highest point density. The dashed line indicates 1 to 1 correspondence.

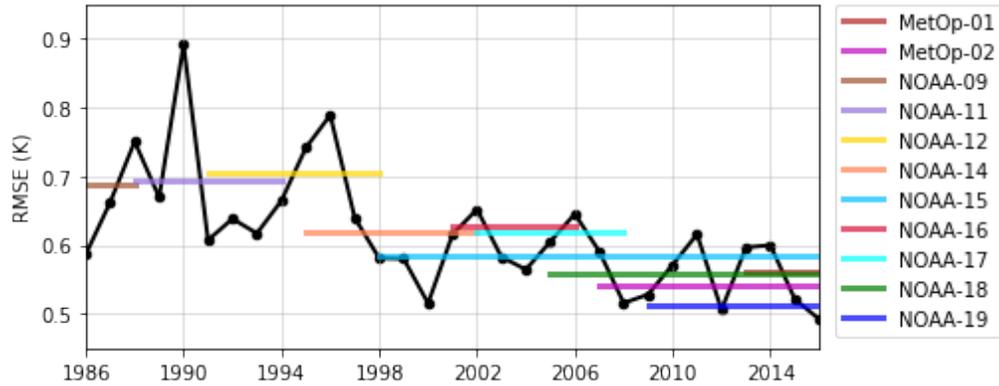


Figure 6. Yearly averaged RMSE for all sensors (black), and each individual sensor (colored) plotted for the lifetime of each sensor

The time of day in which a measurement is made could also impact the calculated RMSE. Larger uncertainties occur in daytime measurements where wind speeds are less than 2ms^{-1} compared to measurements made during the night when wind speeds are greater than 2 m s^{-1} (Wilson et al 2013). Figure 7 shows the average, and standard deviation of the errors at various wind speeds separated by day and night. At low wind speeds, daytime measurements show a positive bias and nighttime measurements show a negative bias relative to the buoy water temperature. Additionally, low wind speeds resulted in a larger variability in the daytime errors relative to the night. These effects become smaller when wind speeds are larger 2 m s^{-1} . 18.5% of all the AVHRR observations with buoys matchups in the region occur at wind speeds less than 2.5 m s^{-1} .

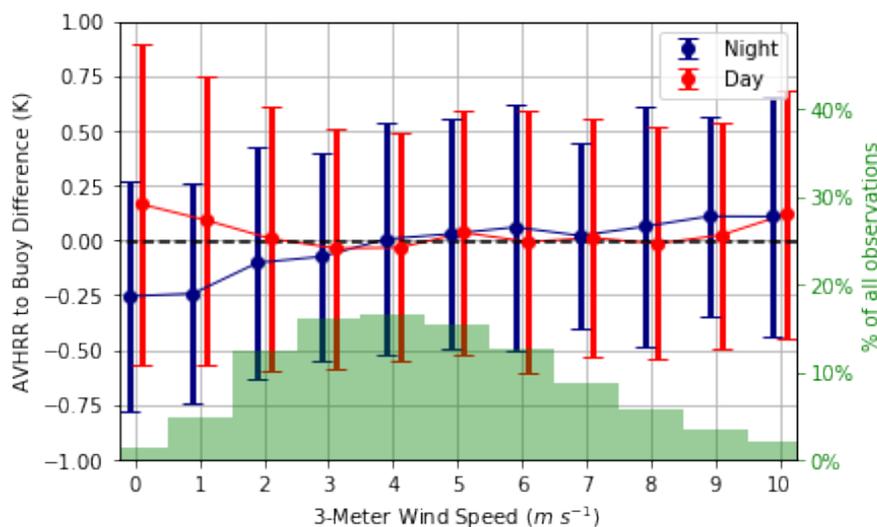


Figure 7. Differences between the AVHRR measured LSWT and the buoy water temperature. Filled circles indicate the mean bias, the error bars indicate the standard deviation of the errors, and the overlaid histogram represents the percent of observations occurring at each wind speed in 1 m/s bins centered on each value. Wind speeds are measured at 3-meters on the buoys.

The NLSST and MCSST equations yielded temperatures that were quite similar (Figure 8). The NLSST was developed to address the nonlinearity in water vapor absorption and at high temperatures. The summertime mean total precipitable water is 29.9 kg m⁻² for the study region (data from North American Regional Reanalysis). It could be that typical atmospheric profiles in the study region are not moist enough, nor temperatures warm enough for the nonlinearity to greatly affect LSWT calculations here. Of the 18,680 matchups, 92.75% of the differences were below 0.05 K, and the maximum NLSST to MCSST difference was 0.24 K. The largest positive differences between the MCSST occurred at very high (>295 K), and at very low (<275 K) water temperatures. More commonly, the NLSST showed only a very small negative difference (roughly -0.02 K) at more moderate water temperatures. The differences

between the NLSST and MCSST were not found to vary greatly with year, viewing angle, or satellite platform.

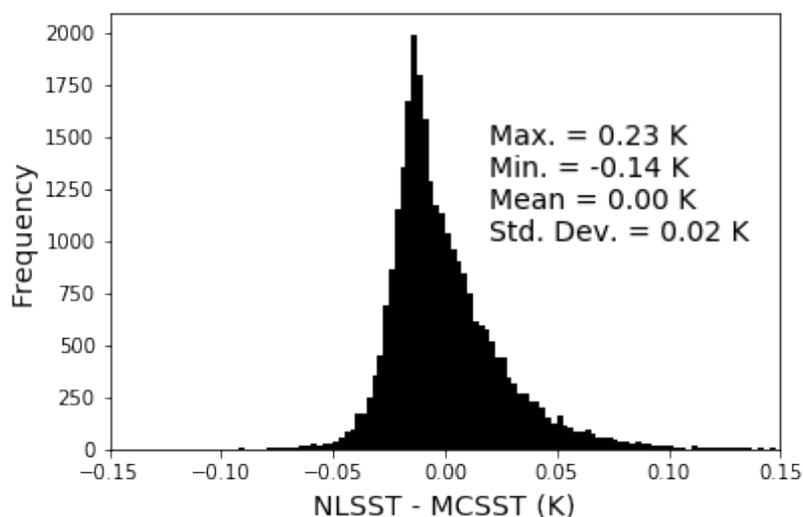


Figure 8. Broad agreement between SST estimates based on the NLSST and MCSST equations applied to AVHRR data from 1986-2016. Bin widths are 0.0025 K. Total n=18,680.

5.4 Diurnal Correction and Compositing

Altogether, the resulting independently calculated diurnal cycles from the buoy data and the AVHRR data show good agreement. The diurnal cycle amplitude in the buoys was expected to be dampened relative to that of skin temperatures observed by AVHRR because the surface is more sensitive to diurnal variability. However, the buoy measurements could perhaps still reproduce the shape of the average diurnal cycle and serve as a reference to ensure realistic results for the timing of the maximum and minimum of the fitted curve for the satellite observations. We characterized the buoy diurnal cycle for each month by calculating the departure from the daily mean for all measurements made between 1986 and 2016. The amplitude of the diurnal cycle changes significantly between months, reaching its maximum amplitude in July (Figure 9). Since the buoys are removed from the lakes during the winter

months, mean diurnal cycles are only shown when there were at least four years where the buoy was present for the entire month.

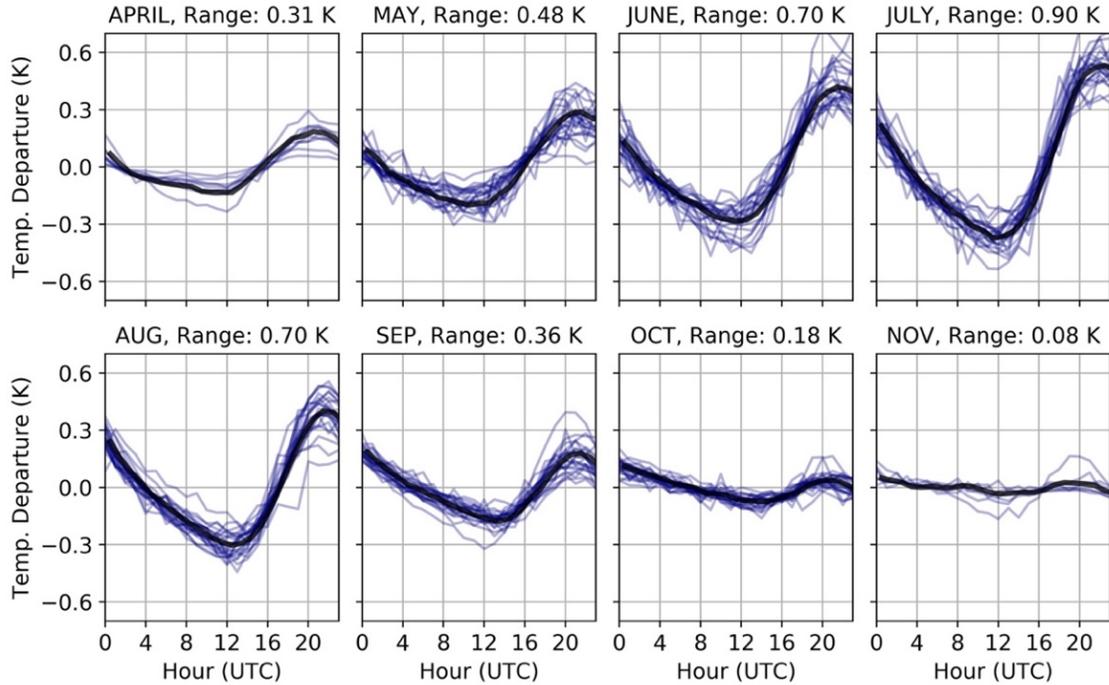


Figure 9. Average diurnal cycles found in the buoy observations for the Laurentian Great Lakes plotted as a function of departure from the daily average temperature. The blue lines indicate the mean for individual years for the specified month, while the black line indicates the mean across all years. The range indicates the difference between the maximum and minimum values of the mean across all years.

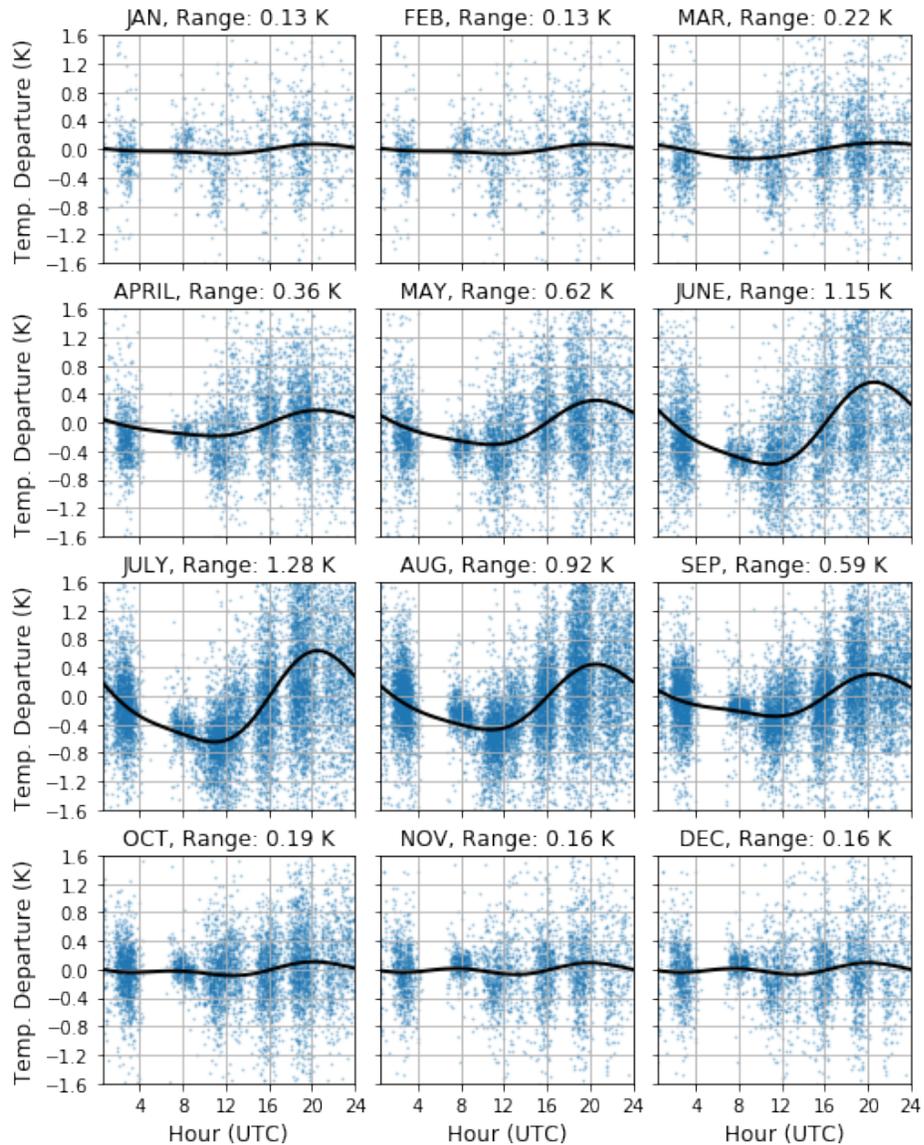


Figure 10. Average diurnal cycles found in the AVHRR observations for April through November. The blue points indicate the individual differences calculated between an observation at the specified time and the observation in the reference window (8-11 UTC). The black line indicates the diurnal cycle function fitted to the centered differences. The range indicates the difference between the maximum and minimum values of diurnal cycle function.

The AVHRR-derived diurnal cycles of LSWT are readily discernable after standardizing, and centering differences to the 8-11 UTC window (Figure 10). There is some agreement between the AVHRR derived cycles, and the buoy derived cycles in the timing between the minima and maxima with a couple of exceptions. The differences between the

maxima and minima are generally only slightly larger for the AVHRR diurnal cycles. Noteworthy exceptions occur for June and July which are 0.45 K and 0.38 K larger. There were not enough differences to fit diurnal cycles for the more cloudy months, so the observations for November and December were combined. The same was done for January and February. This could increase the error in our process, but the minimal diurnal variability in the colder months makes large errors unlikely. Nonetheless, this should be noted when calculating trends for wintertime LSWT.

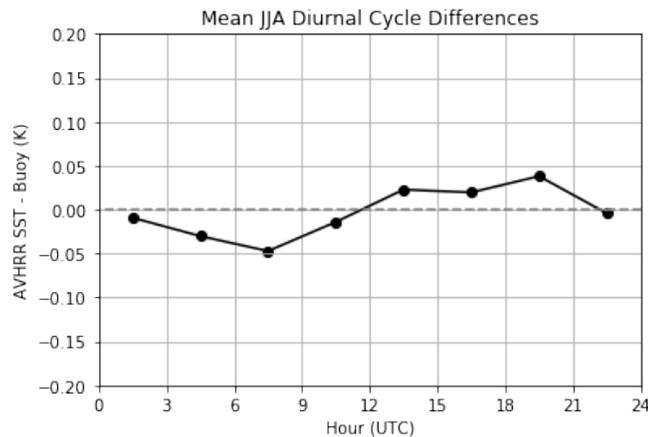


Figure 11. Mean differences between AVHRR SST and buoy water temperature in June, July and August. Each filled circle represents the mean error for the 3 hour window centered on the circle.

To examine the differences between the amplitude of the buoy and AVHRR diurnal cycles, mean summertime differences are calculated at different times of day (Figure 11) between the AVHRR LSWT and buoy water temperature measurements. AVHRR LSWT exhibits a positive bias during the day, and a negative bias during the night relative to the buoy measurements. When comparing the AVHRR SST and buoy diurnal cycles, the former has a slightly large amplitude. The mean JJA differences (Figure 11) are relatively small compared

to the AVHRR diurnal cycle amplitude because they are occurring on-top of the observed buoy diurnal variability.

The fitted diurnal cycle functions are then used to create daily composites as is done in Foster and Heidinger (2013). Individual SST measurements are used to estimate the hourly measurements of LSWT throughout the day according to the fitted diurnal cycles and the time the measurement was made. The hourly estimates are then used to compute a daily average. This reduces the total number of images in the time series from 63,470 ascending and descending node composites from all sensors to 11,323 daily images. Daily LSWT composites are only reported where there is at least one clear-sky and ice-free observation of the lake surface. The number of all-sky views, and the number of days with at least one clear-sky and ice-free view of the lake increases throughout the record with the number of operational AVHRR sensors (Figure 12).

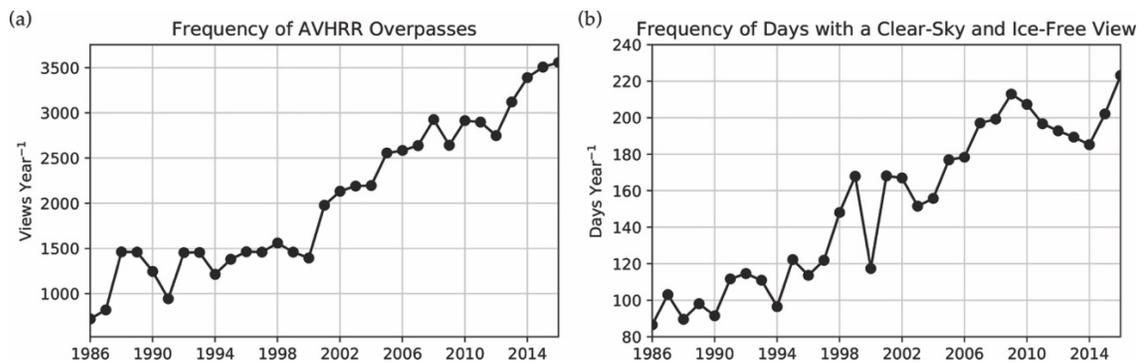


Figure 12. (a) Number of all-sky views of Lake Ontario, and (b) the number of calculated daily composites

5.5 Gap Filling and Cross Validation

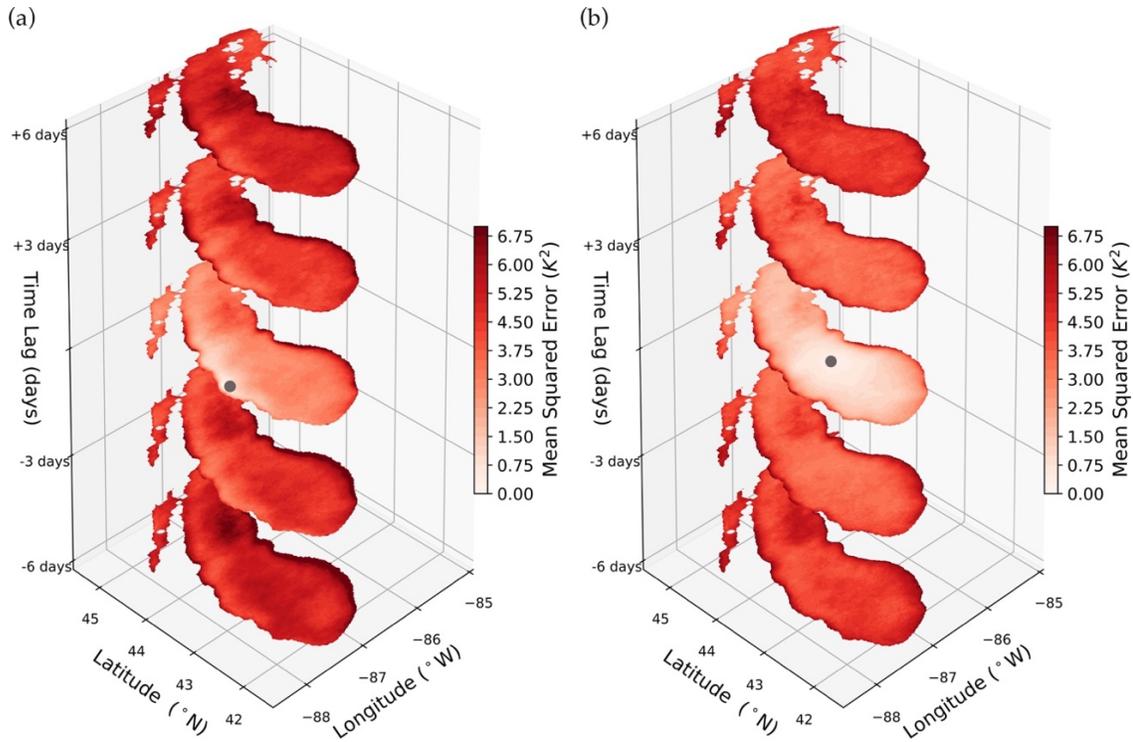


Figure 13. The calculated mean squared errors between a location (a) near the western shoreline of Lake Michigan, and (b) towards the middle of Lake Michigan plotted for time lags of -6,-3,0,3, and 6 days in July. The grey point indicates the location from which mean squared errors from all other locations are calculated to.

An example of the mean squared errors used to calculate the weights in equation 6 are shown in Figure 13a for a location close to the western shoreline of Lake Michigan and in Figure 13b for a location in the middle of the lake for the month of July. As expected, differences typically increase as the spatial and temporal separation between two observations increases. However, there are many cases where distant observations have lower errors than nearby ones. This scenario commonly arises along shorelines, where points just offshore have much lower errors than those adjacent to the shoreline.

During the month of February, deseasonalized LSWT is much more homogenous than during the warmer months, resulting in spatiotemporal variation being relatively low. Despite

there being a greater reliance on interpolation to fill missing values in the colder months, a given observation is likely to have only a small difference to its neighbors in across space and time. (Figure 14).

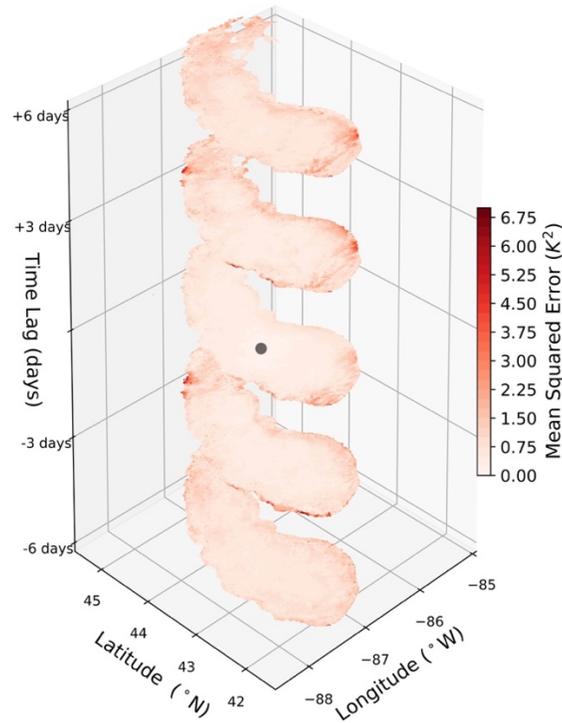


Figure 14. Same as Figure 13, but for the month of February for a mid-lake point. For some locations, values are not computed due to an infrequency of clear-sky and ice-free observations of the lake during the winter.

Since clouds often result in multi-day gaps spanning a large portion of the lake surface, a weighted average of local anomalies is not always optimal. Here, we have performed a simple cross validation to illustrate how a combination of local averaging and LOESS can give improved estimates of LSWT under cloudy conditions (Table 3). The reported errors show that using LOESS alone (RMSE = 1.26K) is suboptimal compared to selectively inserting locally weighted averages prior to using LOESS (RMSE = 1.10 K). All of these methods demonstrate low bias ($\leq 0.1K$), and a moderate correlation between the estimated and observed residuals ($\rho \geq 0.67$). The local averages perform relatively poorly where there are very few observations

in a local neighborhood, and where heterogeneity is large in the neighboring observations. No metric is used to assess the confidence of individual LSWT measurements, so the local averages can tend to propagate erroneous values when they are present. Thus, the aim is use local averages where there is small chance they will propagate erroneous values, and improve the resulting LOESS fit by only inserting local averages into the time series where there the local neighborhood is spatially and temporally homogenous.

Table 3. Results from the cross-validation. Root mean squared error (RMSE), mean bias, and Pearson correlation coefficient (ρ) are calculated between the three different methods used: a locally weighted average (LWA), LOESS, and a combination of selective LWA and LOESS. $n=33,593$.

Gap Filling Method	RMSE	Bias	Correlation Coefficient (ρ)
LWA	1.57 K	0.10 K	0.67
LOESS	1.26 K	0.04 K	0.75
LOESS + Selective LWA	1.10 K	0.05 K	0.86

6 Analysis of the Record

6.1 Mean Intra-annual Variation

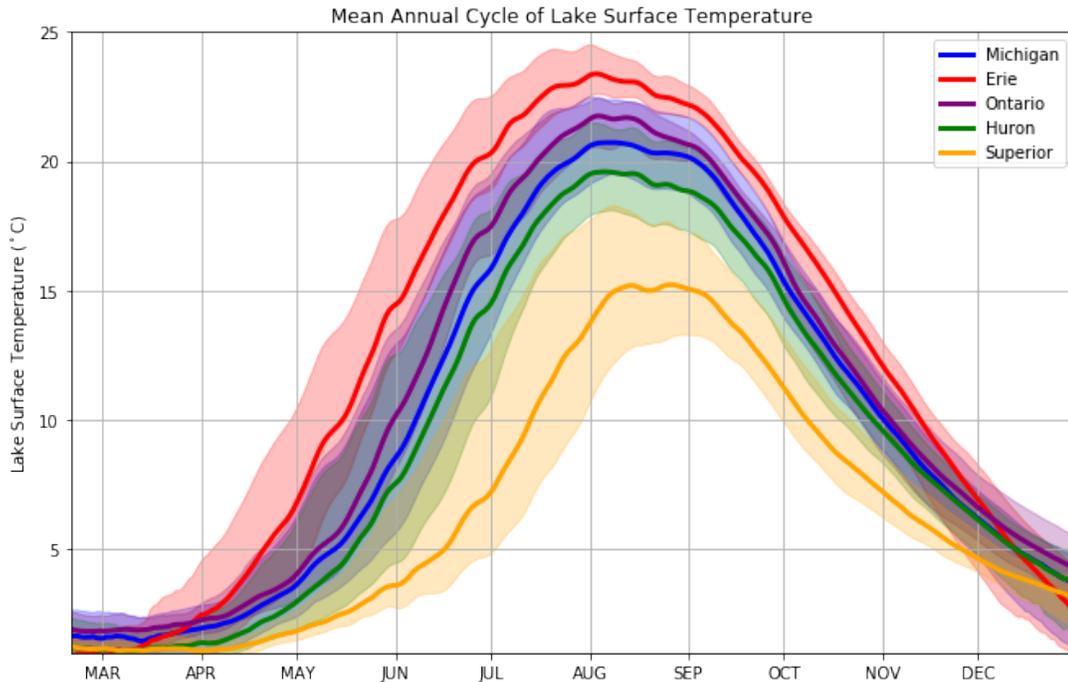


Figure 15. Climatology of the mean lake temperature, and the range of temperatures in each of the five lakes. The thick line represents the 31-year climatology of lake-wide mean temperature. The top and bottom of the shaded areas represent the climatology of the 95th and 5th percentile of the lake-surface temperature distribution across each lake. Julian days between 0 and 50 are not shown due to frequent ice cover.

A climatology of lake surface temperature is created by calculating the lake-wide mean temperature for each Julian day for the entire 31-year record. The five lakes demonstrate a very wide range of temperatures across space for the average year (Figure 15). The shallowest lake, Lake Erie, warms much faster, reaches a higher maximum temperature, and cools more quickly than the other 4 lakes. In contrast, Lake Superior, the largest lake in this study, warms and cools more slowly, and reaches a much lower maximum temperature and cools more slowly

compared to the other lakes. Additionally, Lake Superior reaches its maximum lake-wide mean temperature much later in the year compared to the other four lakes.

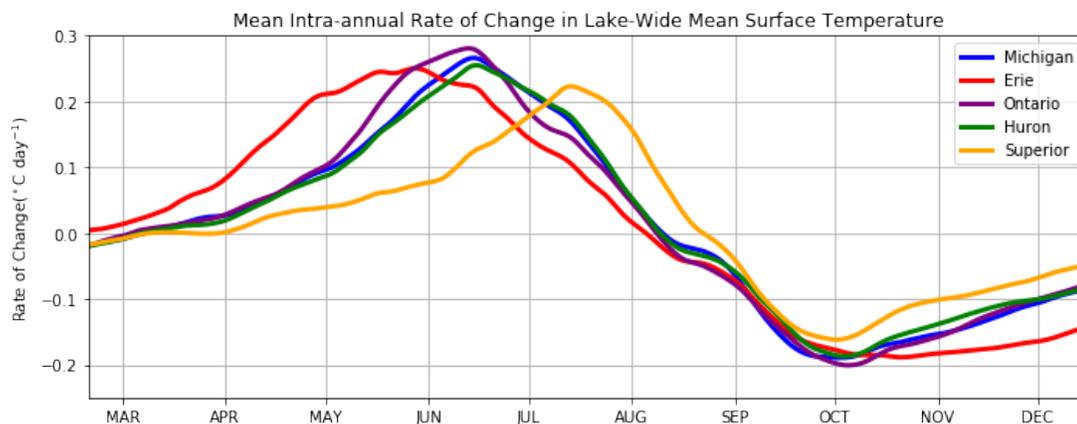


Figure 16. Climatological mean rate of change of lake-wide mean surface temperature (1986-2016). The rate of change is calculated by a 30-day moving linear regression.

The rate of change of the mean lake surface temperature is calculated by a moving 30-day linear regression (Figure 16). The largest lakes with the deepest bathymetry, reach peak mean intra-annual warming much later in the year, than Lake Erie (the smallest and shallowest). On average, all lakes begin warming between February and March, begin cooling between mid-August and mid-September, and reach peak mean intra-annual cooling in October

To further investigate the differences in the mean intra-annual variation of the lake surface temperature, the intra-annual rate of change of lake surface temperature is separated for the shallowest and deepest 10% of each lake (Figure 17). There are pronounced differences between the two areas of these lakes. The deepest areas reach slightly larger peak warming rates that occur later in the year compared to the shallower areas. This occurs despite the shallowest areas reaching higher temperatures earlier in the year compared to the deeper areas.

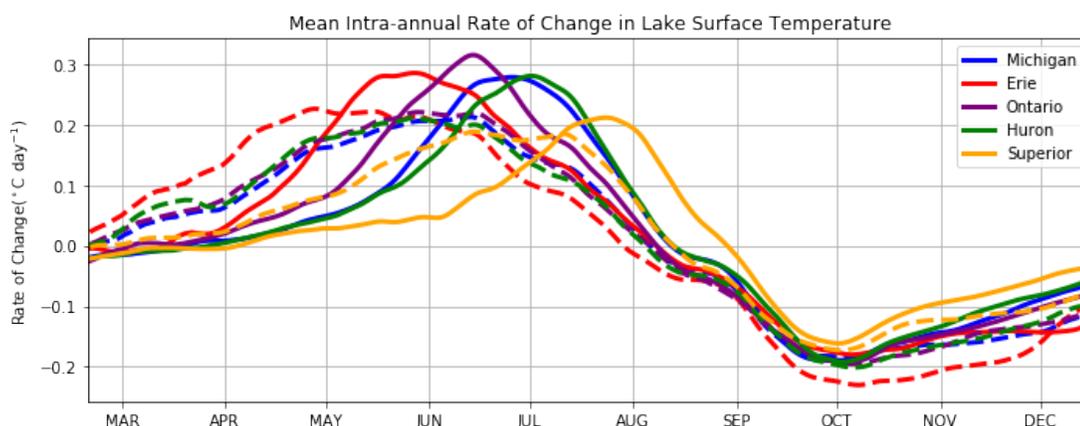


Figure 17. Climatological (1986-2016) mean intra-annual rate of change of the shallowest 10% (dashed line), and the deepest 10% (solid line) of each lake. The rates of change are calculated by a 30-day moving linear regression.

The annual cycle of lake surface temperature is also accompanied by a seasonal variation in surface temperature heterogeneity (Figure 18). The timing of the maximum within-lake thermal heterogeneity varies significantly between lakes. The deepest lakes, reach peak heterogeneity later in the year than the shallowest lakes. The seasonal cycles of heterogeneity in Lake Ontario, and Lake Erie are bimodal and reach a second, smaller local maximum in late autumn. A similar analysis was performed using the range (maximum-minimum) of surface temperature and the same pattern as the one shown in Figure 18 was found for all lakes.

The climatology of lake surface temperature also indicates a strong seasonality to the magnitude of interannual variability of mean lake surface temperature (Figure 19). The lake-wide mean temperature is variable between years and the maximum interannual variation typically occurs in spring and early summer. This maximum in interannual variation matches very closely in date to the maximum in mean lake heterogeneity for all lakes except Erie. All lakes show relatively low interannual variability during the fall, winter and early spring.

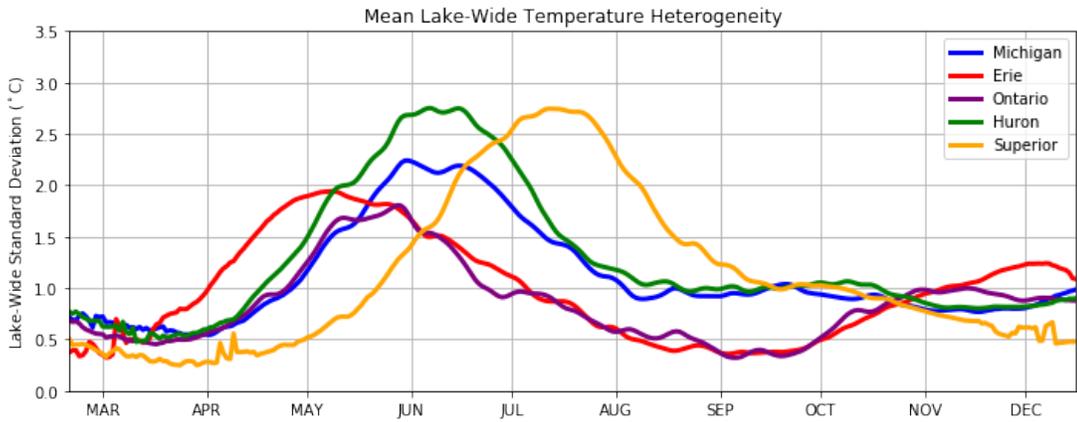


Figure 18. Climatological standard deviation of temperature across the lake surface for a given day of the year.

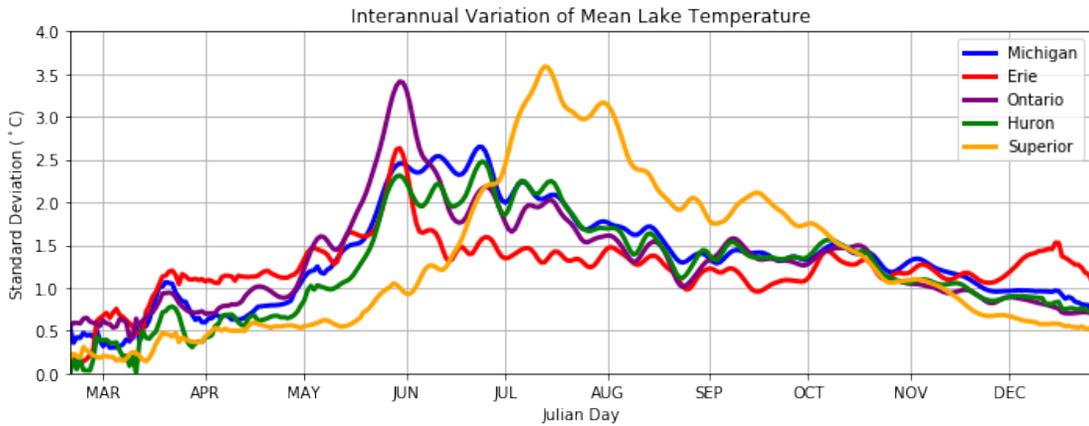


Figure 19. Interannual variation of mean lake temperature. The values of indicated by each line indicate the standard deviation of 31 mean lake temperatures (1 per year) for each day of the year for five lakes.

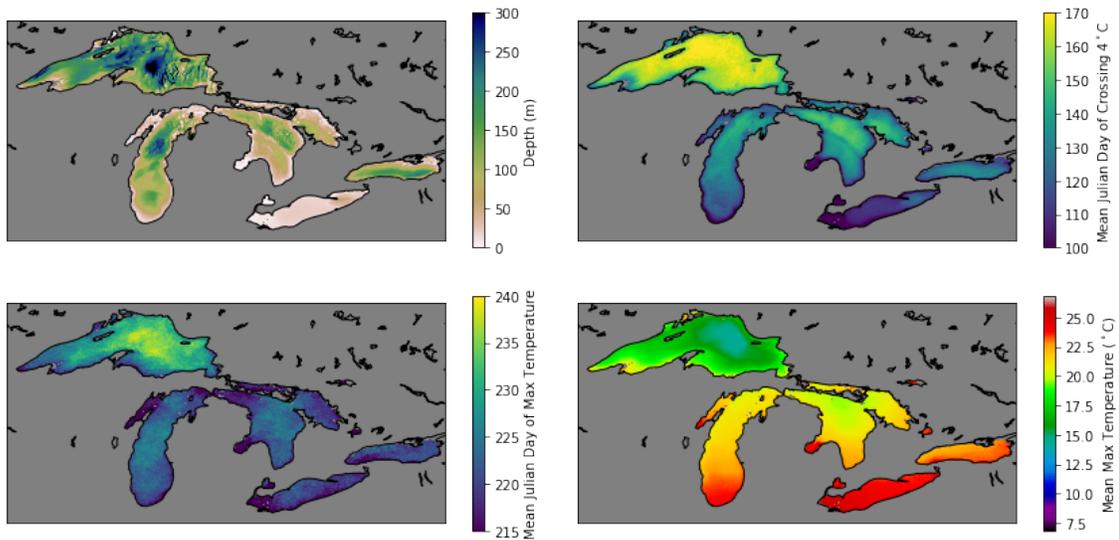


Figure 20. Bathymetric and thermal properties of the Laurentian Great Lakes. (a) Bathymetry, with increasing positive values indicating deeper locations (b) Mean Julian day of crossing the 4°C threshold (c) Mean Julian day of reaching 31-year mean maximum temperature.

A summary of the bathymetric and mean intra-annual thermal properties of the five largest lakes in this region is shown in Figure 20. We take the date of crossing 4 degree celsius threshold as a proxy for the timing of stratification as is done previously (Woolway and Merchant 2018). The date at which stratification occurs (Figure 20b) varies with the depth of the lake in almost all cases. This relationship is most apparent in Lake Huron. The relationship to bathymetry is still present in the date of reaching maximum temperature (Figure 20c), but is less apparent compared to the date of stratification. Lake Superior has marked differences from the other lakes in a number of ways. Lake Superior exhibits a larger range of dates of maximum temperature, which could be a function of its wider range in bathymetry. Lake Superior also reaches stratification much later than the other four, the dates of maximum temperature are typically later, and temperatures much cooler than the other lakes which could be attributed to its slightly higher latitude, and greater mean depth. Larger maximum

temperatures are typically reached in the southernmost areas of these lakes, and the shallower bay areas such as Green Bay, Saginaw Bay, and western Lake Erie.

6.2 Long-Term Trends

Trends in July-August-September (JAS) mean surface temperatures are found by calculating linear regression slopes on the yearly mean temperature for these months (as done in Schneider and Hook 2011; O'Reilly et al. 2015; Mason et al. 2016). Lake Superior shows the largest warming rates of roughly 0.8 degrees per decade in the deeper areas of the lake (Figure 21). Large heterogeneity in warming rates is shown both within each lake individually and between lakes.

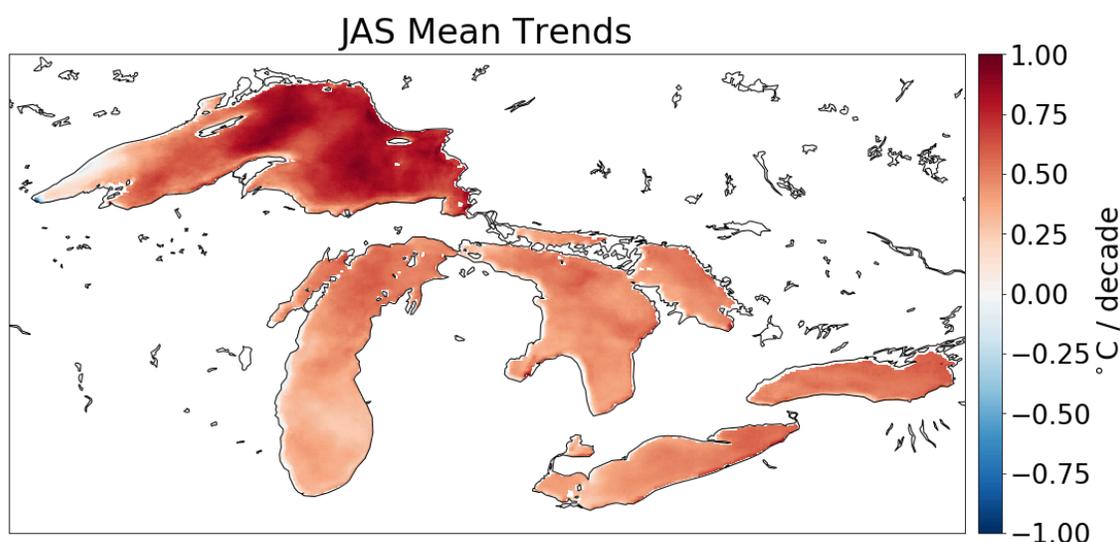


Figure 21. 31-year (1986-2016) linear regression slopes for mean July-August-September (JAS) surface temperatures plotted in degrees Celsius per decade.

These warming rates can be calculated separately on a monthly basis to further examine this apparent heterogeneity. There are large differences in trends depending on the portion of the year examined (Figure 22). The variation of trends across months in many cases is comparable to, or larger, than the spatial variation of trends within months. While July, August and September do indicate warming that is apparent in Figure 21 there is large variability in

these warming rates between the three months ranging from 0.8 degrees Celsius per decade to no trend at all. For Lake Michigan a JAS average does not capture the warming that is occurring in the deeper areas of the lakes in June and fails to capture the largest warming signal observed in October.

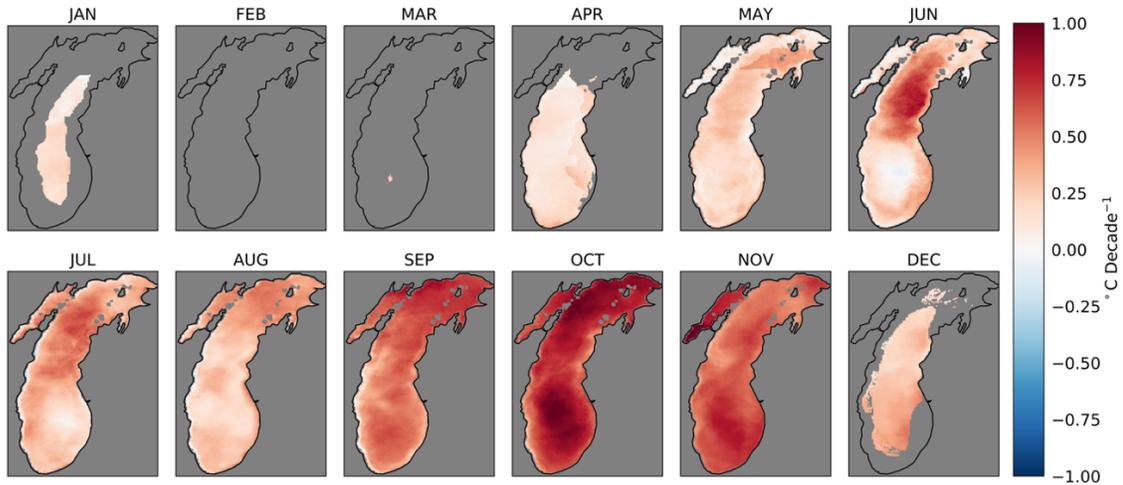


Figure 22. Trends in monthly mean temperature for Lake Michigan. Trends are only shown for months and locations where more than 28 years are ice-free.

This pattern observed in lake Michigan, with the largest warming occurring in the early autumn, also applies to the other large lakes in this region (Figure 23). Across all lakes, September and October are the fastest warming months, and far exceed the trends in late spring and early summer.

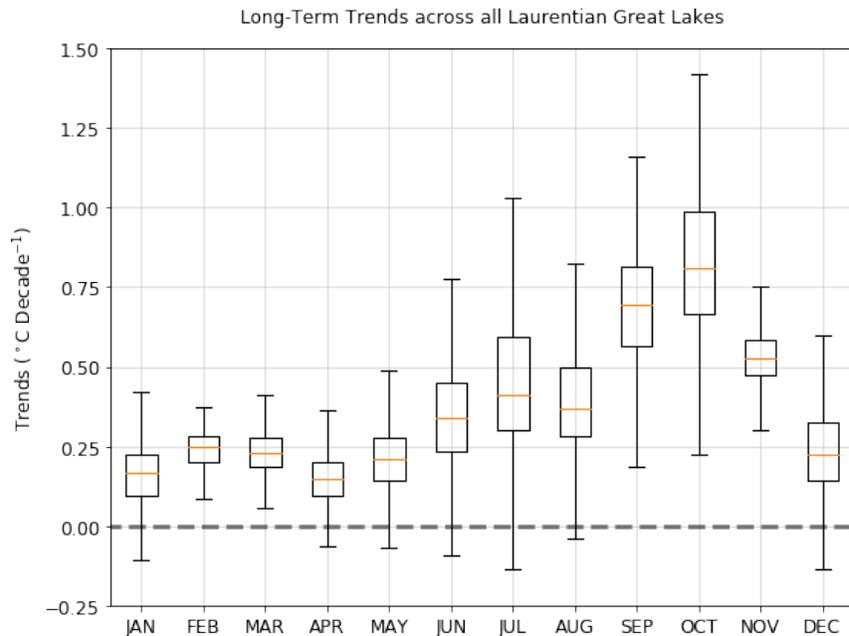


Figure 23. Surface temperature trends (1986-2016) across all Laurentian Great Lakes for each month of the year. The orange line indicates the median trend, box edges indicate the 25th and 75th percentile, and the whiskers indicate the last values within one interquartile range of the 25th and 75th percentiles.

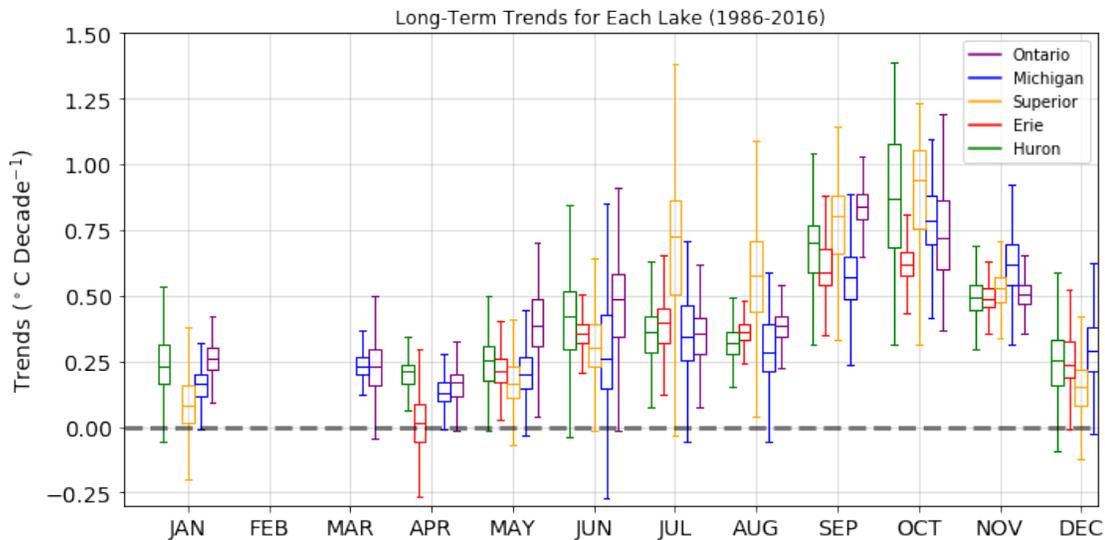


Figure 24. Same as Figure 23, but each box is shown for individual lakes.

The 31-year trends are further separated out for the five largest lakes in the domain to prevent overrepresentation of larger lakes (Figure 24). In many of the months, there are inter-lake differences between in warming trends. Most notably is Lake Superior in July and August,

where the 25th percentile of trends is higher than the 75th percentile in the other four lakes. However, the general pattern of largest warming in October, and November still holds for all lakes individually.

Pearson correlation coefficients between the warming rates at every location on each of the lakes to the climatological mean monthly temperature, and the bathymetry are calculated. There is a general negative association between warming and mean temperature (Figure 25a). This illustrates that, with some exceptions, much of the warming is occurring in areas of the lakes that have lower mean temperatures. A related pattern is found when comparing to the bathymetry of these lakes, with the deeper areas warming more quickly than the shallower areas (Figure 25b).

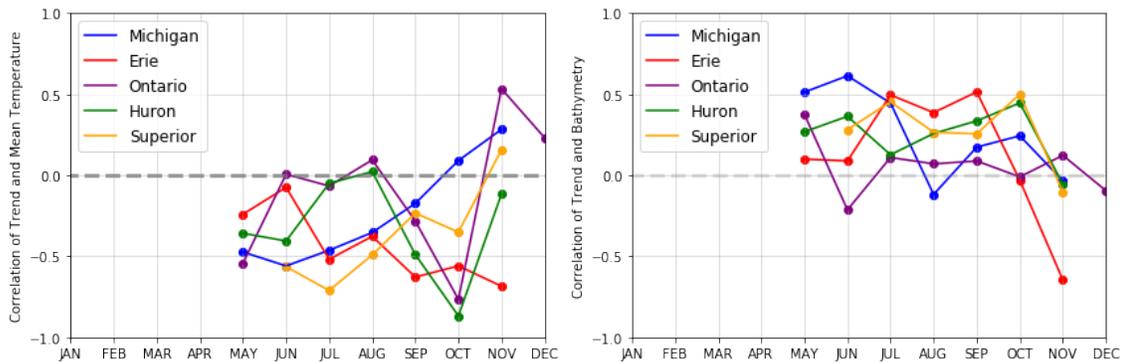


Figure 25. Pearson correlation coefficient between the monthly 1986-2016 trends, and the (a) climatological mean monthly temperature and (b) bathymetry.

It has been previously suggested that spatial variability in warming rates during the JAS period could be attributable to earlier spring stratification, particularly in deeper areas of these lakes. Here, we find long-term trends in our stratification proxy. Linear regression slopes are computed in stratification date at all locations across all five lakes. In all locations, but for a few exceptions in the shallowest areas (<50 meter depth), the lake surface is becoming

stratified earlier in the year. The largest trends are occurring the in deepest areas of the lakes, with some variability between lakes. (Figure 26).

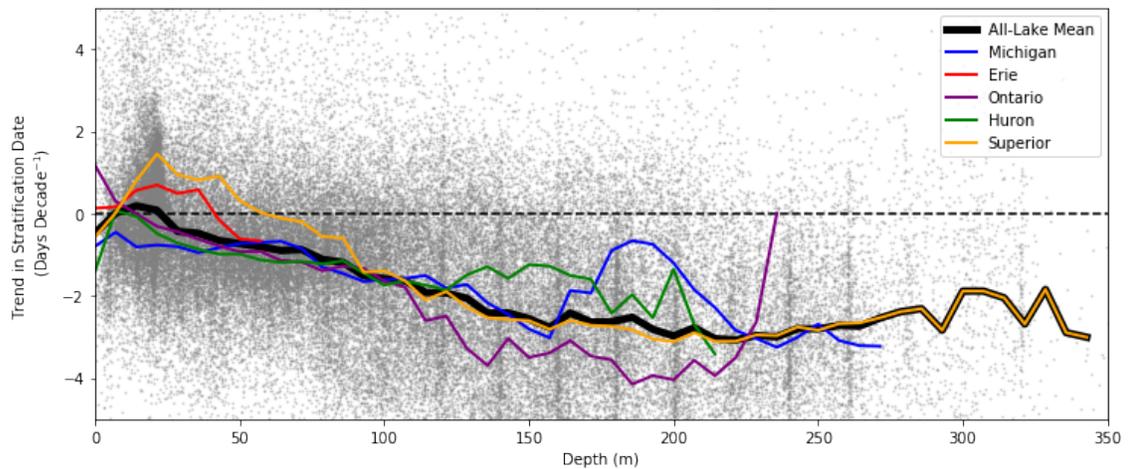


Figure 26. Trends in stratification date as a function of lake depth. Values indicated by the black and colored lines represent means calculated in 7.14 meter intervals. A negative trend implies stratification is occurring earlier in the year with respect to the mean date.

6.3 Spatiotemporal Variability

To further investigate the spatiotemporal variability in the Laurentian Great Lakes, Empirical Orthogonal Functions (EOFs) are found. The lake surface temperature time series is first deseasonalized by subtracting out the 31 year climatology. The climatology is found in the same way as the values plotted in Figure 15, except it is performed separately for every grid cell in the domain. A linear trend in the 31 year time series is then removed, and the EOFs are found to the time series using only days where an individual lake is completely ice-free.

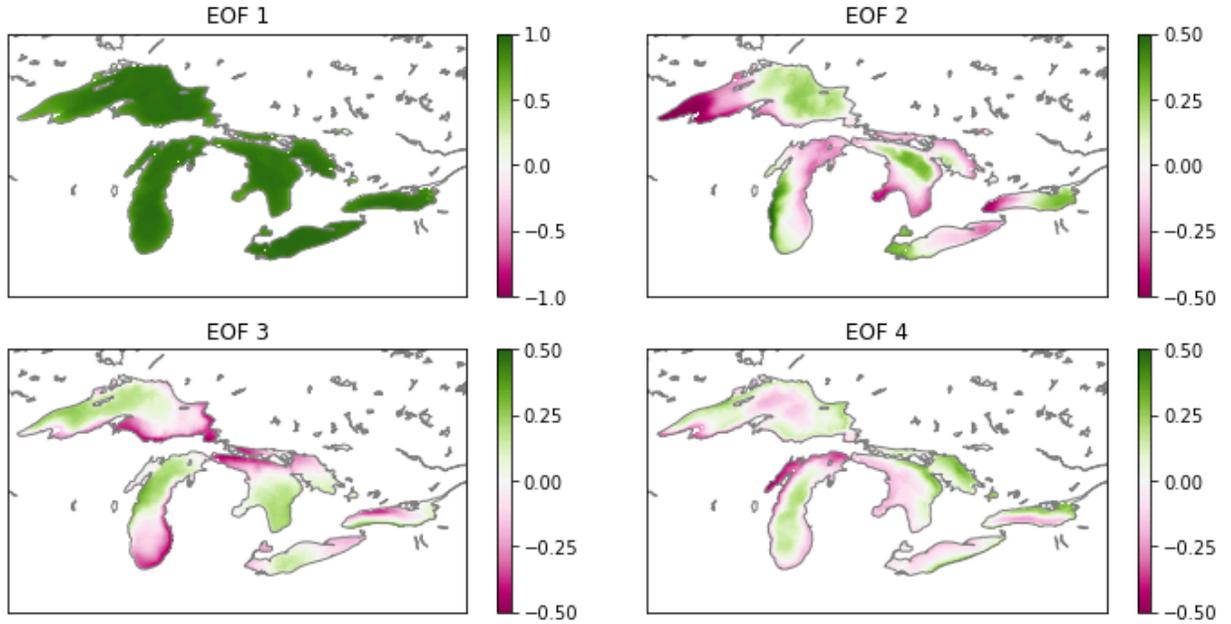


Figure 27. The first four empirical orthogonal functions (EOFs) calculated for the deseasonalized and detrended surface temperature time series. The EOFs are plotted as a function of correlation coefficient between their corresponding principle component, and the surface temperature time series at each grid point. Each EOF is calculated independently for each lake.

The EOFs reveal the statistical modes that explain the largest amount of variance in the deseasonalized and detrended time series. The first four modes are found separately for each lake (Figure 27). A large majority of the variance in these time series are explained by the lakes being uniformly warmer or cooler compared to climatology across the entire lake surface (Table 4). This pattern is true for the first EOF mode across all lakes in this domain. All other modes explain less than 6% of the variance.

Table 4. Percent of variance in the surface temperature time series that is explained by each of the EOF modes shown with the associated errors in the eigenvalues expressed as a percent.

Mode	Michigan	Erie	Ontario	Huron	Superior
EOF 1	$78.81 \pm 0.013 \%$	$85.05 \pm 0.013 \%$	$77.25 \pm 0.011 \%$	$78.15 \pm 0.014 \%$	$78.97 \pm 0.017 \%$
EOF 2	$4.03 \pm 0.001 \%$	$3.53 \pm 0.001 \%$	$5.85 \pm 0.001 \%$	$3.80 \pm 0.001 \%$	$5.82 \pm 0.001 \%$
EOF 3	$3.72 \pm 0.001 \%$	$1.37 \pm <0.001 \%$	$3.01 \pm <0.001 \%$	$3.13 \pm 0.001 \%$	$3.07 \pm 0.001 \%$
EOF 4	$2.67 \pm <0.001 \%$	$1.21 \pm <0.001 \%$	$2.26 \pm <0.001 \%$	$1.74 \pm <0.001 \%$	$1.27 \pm <0.001 \%$

The other modes appear to share some common attributes. The second mode for Lake Huron and the fourth mode for Lake Michigan show a strong correlation to bathymetry. Other modes show moderate correlation in areas of frequent coastal upwelling, like the second mode of Lake Michigan and the third mode of Lake Ontario. The approximate errors of each of the EOF modes are calculated by the method of North et al. (1982). These errors, shown in Table 3, suggest that all of the first four EOFs for all of the lakes in this domain are distinct and do not overlap.

6.4 Upwelling

In order to document the frequent coastal upwelling events in the Laurentian Great Lakes, we have developed an upwelling identification algorithm from the daily lake surface temperature time series. This algorithm largely builds from the algorithm developed in Plattner et al. (2006). The criteria used to identify upwelling during the stratified season are described in Table 4. Given that all three tests in Table 5 are satisfied for a grid cell, and the lake is ice-free, then the grid cell is flagged as having been effected by upwelling. To illustrate how these three tests and their corresponding thresholds behave, two time series are extracted from different locations in Lake Michigan for 2015 (Figure 28). The three tests described in Table 5 are performed on both of these time series and the results are plotted for the inshore location in Figure 29 and an offshore location in Figure 30.

Table 5. The tests used to determine areas effected by large upwelling events, their physical reasoning for inclusion in this algorithm, and their quantitative implementation.

Test Name	Qualitative Reasoning	Quantitative Implementation	Notes
Cooling	When upwelling begins, the temperature at the surface will decrease as cooler water from the bottom of the lake is transported upward towards the surface.	Linear regression slopes are calculated for a 14-day moving window for each grid cell, and for the lake-wide mean. This test is passed for the proceeding 30 days when the regression slopes for a grid cell are 0.25°C/day lower than the lake-wide mean trend.	A difference is taken between the local change and the lake-wide change to distinguish temperature changes due to local upwelling from lake-wide changes in temperature.
Contrast	When looking spatially across the lake surface, upwelling areas have significantly lower temperatures than other areas.	At each grid cell, the median temperature from the surrounding in the surrounding 2°latitude/longitude region is calculated. This test is passed when the temperature of the grid cell is 2.5° lower than the calculated median.	This test has been adapted from the classification presented in Plattner et al (2006).
Coastal	Upwelling events are often confined or contiguous to the coastline, where there is a vertical boundary.	A Pearson correlation coefficient is computed between the temperatures in the surrounding 2° latitude/longitude region, and the distance from the shoreline in the same region. This test is passed when the correlation is <0.5.	This test is included to filter out cases where the inner, deeper areas of the lakes warm more slowly that the shallower areas during the spring.

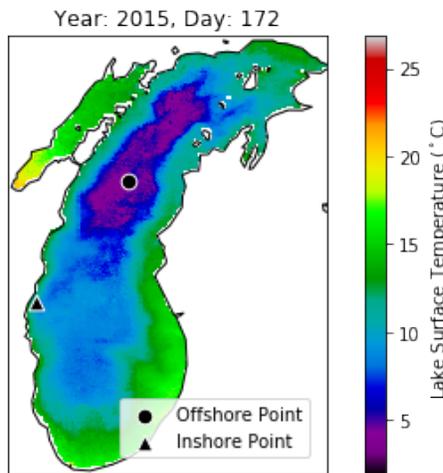


Figure 28. Surface temperature image from Jun 21 2015. Shown are the two points, one inshore, and one offshore, that are referenced in Figure 30 and Figure 31

The time series analysis for the inshore point (Figure 29) indicates four distinct upwelling events that occur between mid-July and mid-November. These events correspond to local minima in the temperature time series, large negative differences compared to the regional median, and were preceded by abrupt cooling indicated by strong negative linear regression slopes compared to the lake-wide trend. The timing and duration of these inshore events are almost entirely specified by the Contrast test in Figure 29c where the local temperature is much lower than the regional median temperature. One event at a Julian day of 260 that was flagged by the Contrast test was not classified as upwelling because it did not meet the requirement set by the Cooling test.

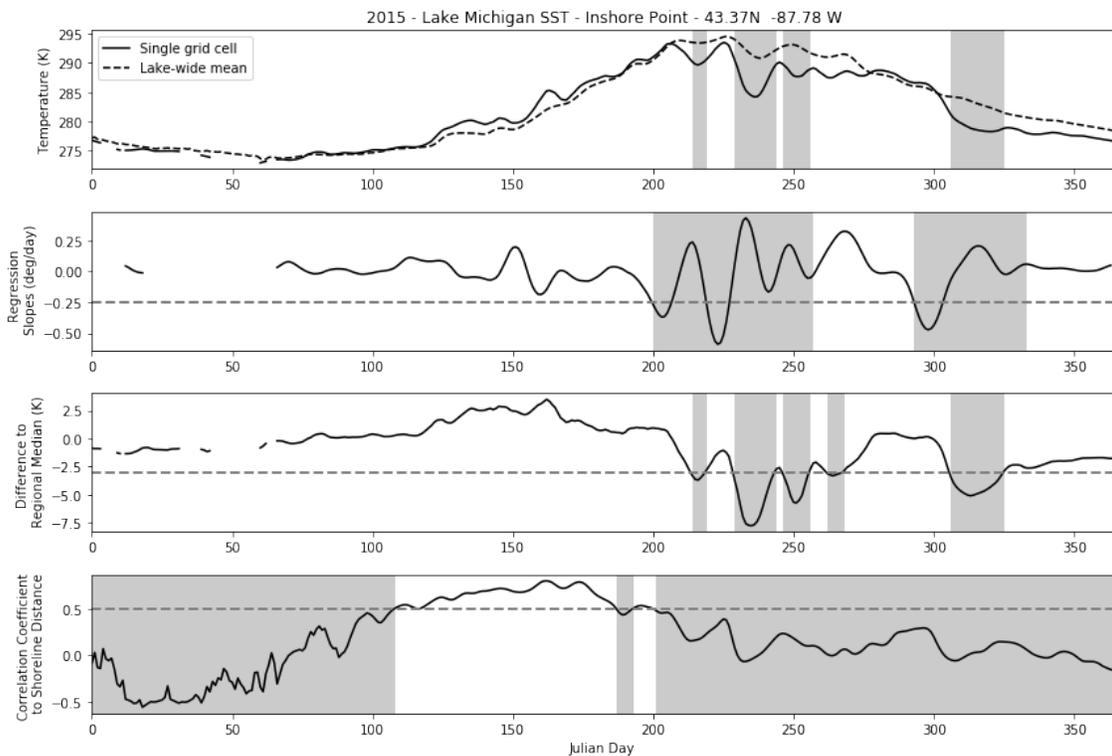


Figure 29. Surface temperature time series and the upwelling tests for the inshore point indicated in Figure 29. The dashed grey lines indicate the thresholds defined in Table 5. The grey shading in (a) indicates where upwelling is occurring defined by this algorithm. The same shading in (b), (c), and (d) are where the individual tests are passed.

There is a roughly 60 day period where the Coastal test threshold was exceeded. This is caused by the shoreline warming more quickly than the innermost areas of the lake. This pattern, of which the lake surface temperature distribution in Figure 29 is indicative, can be seen in the temperature time series in Figure 30a. The selected grid cell along the shoreline had a higher temperature compared to the lake-wide mean which was lower due to the cold inner lake.

While the Cooling test (Figure 29b) and the Coastal test (Figure 29d) have clear physical justification, their inclusion in this algorithm is not appropriately justified by examining a single inshore location. For the inshore case, they provide very little information with the exception of preventing one relatively small event from being flagged as upwelling.

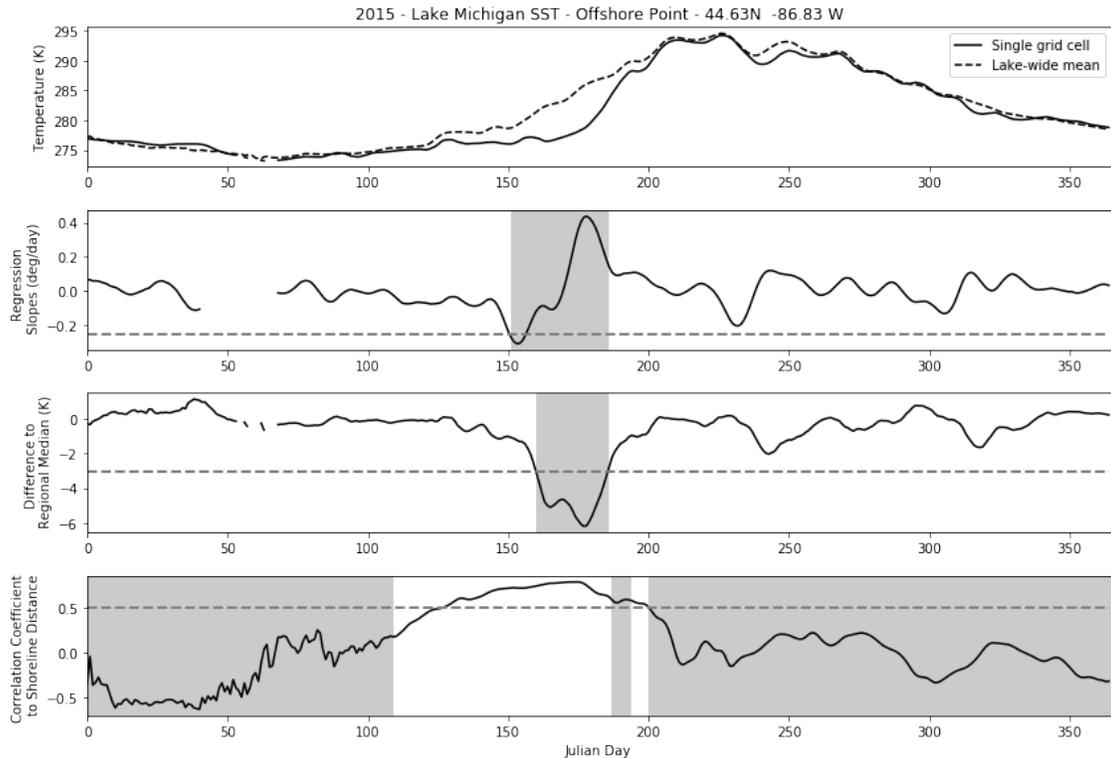


Figure 30. Same as Figure 30 but for the offshore point indicated in Figure 29

The reason for including the Cooling and Coastal tests becomes apparent after examining the offshore time series. It is very uncommon for regions of upwelling to extend offshore into the middle of the lake. However, there are many cases throughout this record where false positives for coastal upwelling occur in the middle of the lake. Figure 30 shows an example of one of these false positives detected by the contrast test. The lake surface temperature image from this false positive is displayed in Figure 28. This pattern occurs when the surface at the shoreline warms more quickly than the surface at the inner portion of the lake.

Figure 30a shows how the lake wide mean warm relatively quickly relative to the local temperature between the Julian days of 130 and 180. This spatial pattern causes the Contrast test to flag a false positive. In this particular case the Coastal test prevented the classification of this false positive due to the low correlation between the regional temperature and shoreline distance, because the cold water was located near center of the lake—which is uncharacteristic of coastal upwelling. The Cooling test was passed because lake-wide mean rate of change was much larger than the local rate of change, the former being driven by the quickly warming coastal areas.

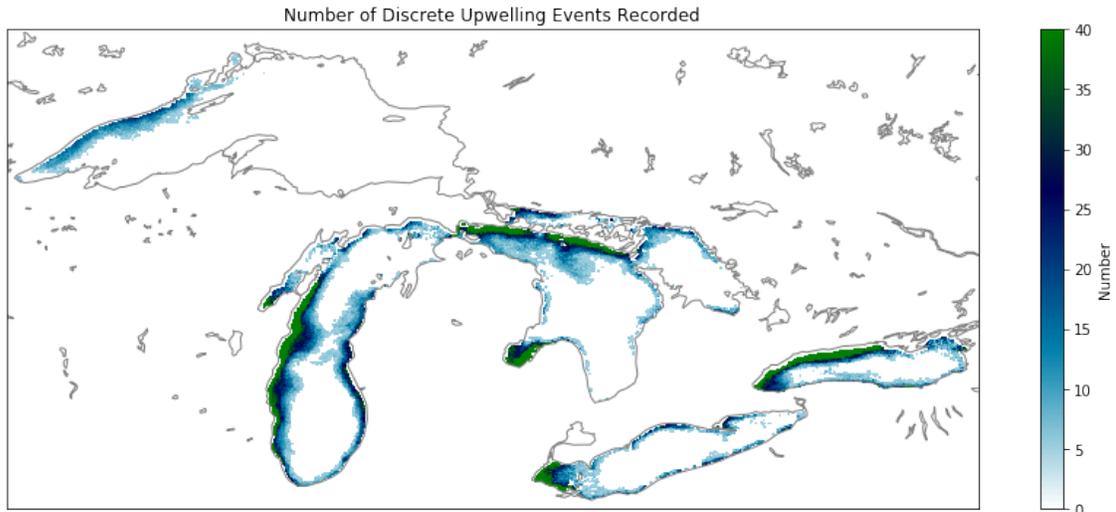


Figure 31. The number of discrete upwelling events identified. Only locations with more than five events are shown.

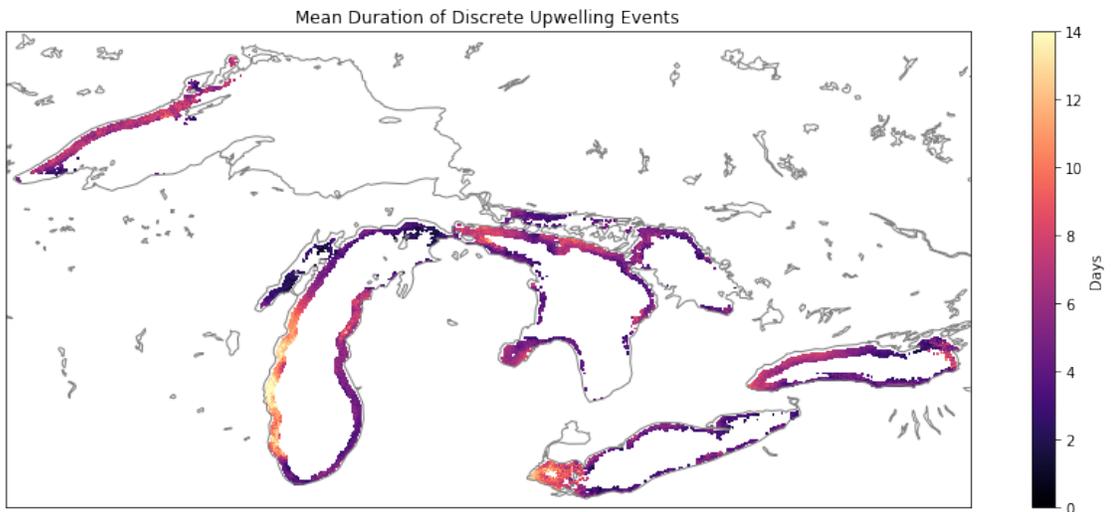


Figure 32. The average duration, in days, of the discrete events shown in Figure 32. Only locations that had more than five documented upwelling events, and are within 5 grid cells of a shoreline are shown.

The tests described above are then applied to all time series across all five lakes on a per lake basis. Discrete upwelling episodes are defined here as consecutive days where a given location passes all the tests. The number of discrete episodes for each location is plotted in Figure 31, and the mean length of these episodes is plotted in Figure 32. From Figure 31, there appears to be several locations where upwelling is relatively frequently detected including the western

and eastern shorelines of Lake Michigan, northern Lake Huron, the northwestern shoreline of Lake Ontario, and the northwestern shoreline of Lake Superior. It also appears that the conditions that make upwelling more frequent in some areas also act to make them last longer.

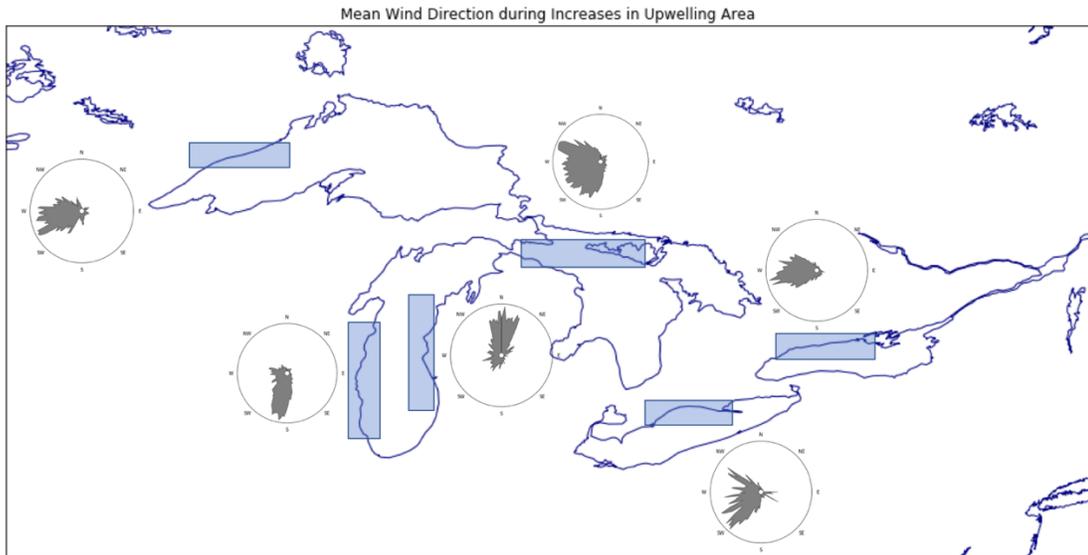


Figure 33. Mean wind direction in each shaded region when an increase in upwelling area occurs in a few selected regions. The wind roses indicate the direction where the wind is coming from.

In an effort to determine whether or not the events that are being documented are actually upwelling, the daily mean 10 meter winds from days where abrupt increases in upwelling area occur are taken from the North American Regional Reanalysis. An abrupt increase in upwelling area is defined as when there is a one-percent increase in the lake surface area classified as upwelling per day in the regions show in Figure 33. The winds are generally parallel to the coastline with the water to the right of the mean wind direction.

7 Discussion

In this work, we have developed a high-resolution, long-term daily surface temperature record for the largest lakes in the Laurentian Great Lakes region. These daily composites were constructed from 11 different AVHRR sensors over the entire 31-year HRPT and FRAC record. LSWT was calculated using coefficients fit on a yearly basis to *in situ* measurements made in this region. There are several potential explanations for differing errors among the sensors (Table 2). Reynolds (1993) noted a negative bias in SST errors associated with the eruption of Mt. Pinatubo in 1991 during the lifetime of NOAA-12. However, these effects were mostly confined to the tropics and the southern hemisphere. Cloud detection plays a prominent role in assessing the accuracy of SST/LSWT measurements. Clouds that are not screened out can contribute significantly to the calculated RMSE. The noise in channel-3 of AVHRR has been documented to be larger in earlier sensors (Dudhia 1989; Warren 1989), and could negatively impact cloud detection. The resulting higher frequency of missed clouds on the earlier sensors could decrease the accuracy of their measurements. The relatively small differences between the NLSST and MCSST are in good agreement with Li et al. (2001) which found a similar result for the Great Lakes.

To account for the navigational errors in the HRPT data, linear adjustments were made to the regrid images in order to align them with a static land-classification field. Overall, the largest impact of these geolocation errors occurs along the coastline where brightness temperature gradients can be large. When these errors are left uncorrected, they can result in abrupt increases and decreases in sea surface temperature for land-adjacent pixels. Since the largest errors occur in the earlier sensor records, the largest impact of these corrections occurs

between 1986 and 1999. For images with extensive cloud cover where no adjustment is made, the impact of poorly georegistered brightness temperatures can be mitigated by discarding values with extremely large and abrupt departures from the climatological mean for a given date.

We have attempted to remove the effects of diurnal variability on our dataset by empirically determining and subtracting the mean diurnal cycle. The differences between the maxima and minima are generally only slightly larger in the AVHRR measurements compared to the buoys. The differences between diurnal cycle amplitudes from the two platforms could be exaggerated because the AVHRR LSWT diurnal cycles are biased towards clear-sky observations, whereas the buoy diurnal cycles represent a wider range of conditions and are more representative of the bulk temperature. This correction performs best where atmospheric and limnological conditions don't stray far from the mean. As illustrated in Figure 7, low wind speeds increase the disparity of daytime and nighttime observations. Under circumstances where the climatological surface wind speed changes, the diurnal correction employed in this record would not account for the increase or decrease in the amplitude of the mean diurnal cycle. The same might occur for a change in mean cloud cover. Furthermore, the diurnal correction we have applied ignores possible variation across space. Lake surface area, and light attenuation (albeit much less so than surface area) have influence over the diurnal temperature range of many lakes (Woolway et al. 2016). It seems plausible that smaller bay areas of these lakes (i.e. Green Bay, Travers Bay, Saginaw Bay,...) could have mean diurnal cycles larger than those calculated using offshore time series due to spatial variation of mean mixed layer depth. Additionally the spatial variation and long-term changes in water clarity experienced in

these lakes (Binding et al. 2015) could perhaps alter the vertical heating profile resulting in changes in mean diurnal variability at the surface.

The examination of the mean AVHRR bias during June, July and August as a function of the time of day (Figure 11) justifies the marginally larger range in the AVHRR LSWT diurnal cycles. The small differences between them are explained by examining the error as a function of the time day, and suggests that diurnal variability is, as expected, larger at the surface. Nonetheless, it is surprising that diurnal variability can be observed in a satellite record in which measurements are taken irregularly and sometimes infrequently. This is made possible by the length of the AVHRR record and the fact that many of the sensors undergo significant orbital drift so that almost all hours of the day are observed.

Gaps due to cloud cover were filled using a combination of locally weighted methods. The calculated weights for one inshore location and one offshore location demonstrate that there are distinct differences in the relationship between a given location on the lake surface, and its neighboring observations. These relationships vary with the reference location, and the zonal, meridional, and temporal separation. Therefore, weighted averages using these empirically determined weights should produce a better result than an unweighted average that ignores these subtleties. It should be noted that the RMSEs reported here may not necessarily be indicative of the errors when interpolating missing values. In practice, interpolation is needed more often in data-sparse regions and times that may not be well represented in our cross-validation. The cross-validation employed here neglects the fact that interpolation is most needed during the winter months where the lake is more homogenous. When the lake is more homogenous, errors from all methods will be lower. Thus, the RMSE values in Table 3 should only be interpreted in a relative sense, to indicate that LOESS supplemented by select

locally weighted averages typically results in smaller errors than using LOESS, or locally weighted averages alone. Due to the low frequency of observations (particularly in the early portion of the record), care should be taken when using the gap-filled product in the winter months where cloud-cover is frequent. When possible, wintertime LSWT features observed in the gap-filled product should be verified in the original daily LSWT composites.

Using this spatially and temporally complete dataset, we have determined the climatology, long-term surface temperature and stratification trends, and some aspects of spatiotemporal variability in surface temperatures of the Laurentian Great Lakes. The lakes analyzed in this work demonstrate varied mean seasonal cycles. The largest and deepest lakes, warm more slowly, reach their maximum temperature and reach peak warming later in the year, and cool more slowly. The maximum warming rates occur during the same time period as the maximum in lake-wide heterogeneity, and maximum interannual variability. The largest values in lake-wide heterogeneity seen are usually during periods of the year where the shallow areas have become stratified and has warmed relatively rapidly, compared to the deepest areas of the lake. This can be identified by comparing the mean date of stratification (Figure 20b) and the climatological difference between the temperature in the shallow and deep areas of the lake (Figure 34). The shallow/deep difference begins to increase sharply around the date in which the shoreline becomes stratified. The difference typically reaches its maximum around the date in which the deeper portions of the lake become stratified and the difference decreases thereafter. Similarly, the negative difference seen in Lake Erie and Lake Ontario in late autumn occur when the shoreline has cooled and the inner region of the lake remains relatively warm. This implies that a large portion of the spatial variability of spring lake surface temperature

(Figure 17) could be associated with the varied bathymetry, and date of stratification in these lakes.

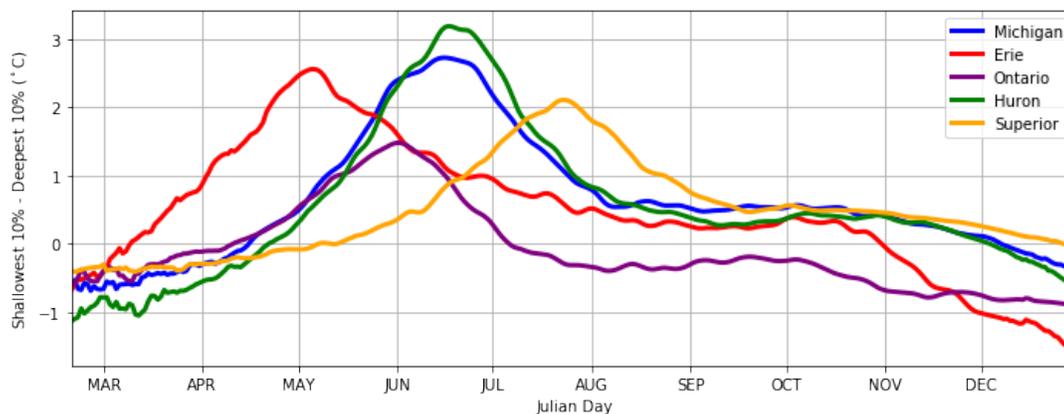


Figure 34. Climatological (1986-2016) mean temperature difference between the shallowest 10% and deepest 10% of each lake.

The long-term trends calculated for summer (July, August and September means), indicate a large degree of variability in warming rates both within and between lakes. Lake Superior summertime surface waters are warming more quickly than the other lakes and demonstrate a larger amount of spatial variability in warming rates. All lakes show indications of slightly enhanced warming of the lake surface over the deeper areas of the lakes with colder surface temperatures (Figure 25).

The values for 1986-2016 JAS warming trends shown here do not agree in magnitude and only moderately agree spatial distribution with those calculated in Mason et al. (2016) for 1994-2013 and in Woolway and Merchant (2018) for 1992-2010. We calculated trends for 1994-2013, and 1992-2010, to compare to our 1986-2016 JAS trend. Across all lakes in this study the 1994-2013 trends are 1.65 times larger than 1986-2016, and the 1992-2010 trends are 2.26 times larger than 1986-2016. The spatial distribution is also affected (Table 6). Lake Superior is relatively unimpacted, but the spatial distribution of the warming trends varies

largely for Lake Huron (only showing a correlation coefficient of 0.29 between 1986-2016 JAS trends and 1995-2013 JAS trends), and Lake Ontario.

Table 6. Pearson correlation coefficient of 1986-2016 trends to the trends calculated for 1995-2013 and 1992-2010 periods using the same AVHRR HRPT dataset.

Lake	Correlation Coefficient to	
	1986-2016 JAS trend	
	JAS 1995-2013	JAS 1992-2010
Michigan	0.58	0.51
Erie	0.69	0.57
Ontario	0.55	0.39
Huron	0.29	0.42
Superior	0.80	0.85

However, when the trends are recomputed so that the time-periods analyzed are the same, there is very good agreement between our JAS trends and those published in the other two works discussed here. This indicates a large sensitivity of magnitude and spatial distribution of JAS warming trends to the start and end dates of the record and the length of the record. The fact that the trends calculated in Figure 35b and Figure 35c match those in Mason et al. (2016), and Woolway and Merchant (2018) very well indicates little to no impact of using a different observation platform (whether it's AVHRR or ATSR) and gap-filling method (whether it's the method used here or an EOF based reconstruction).

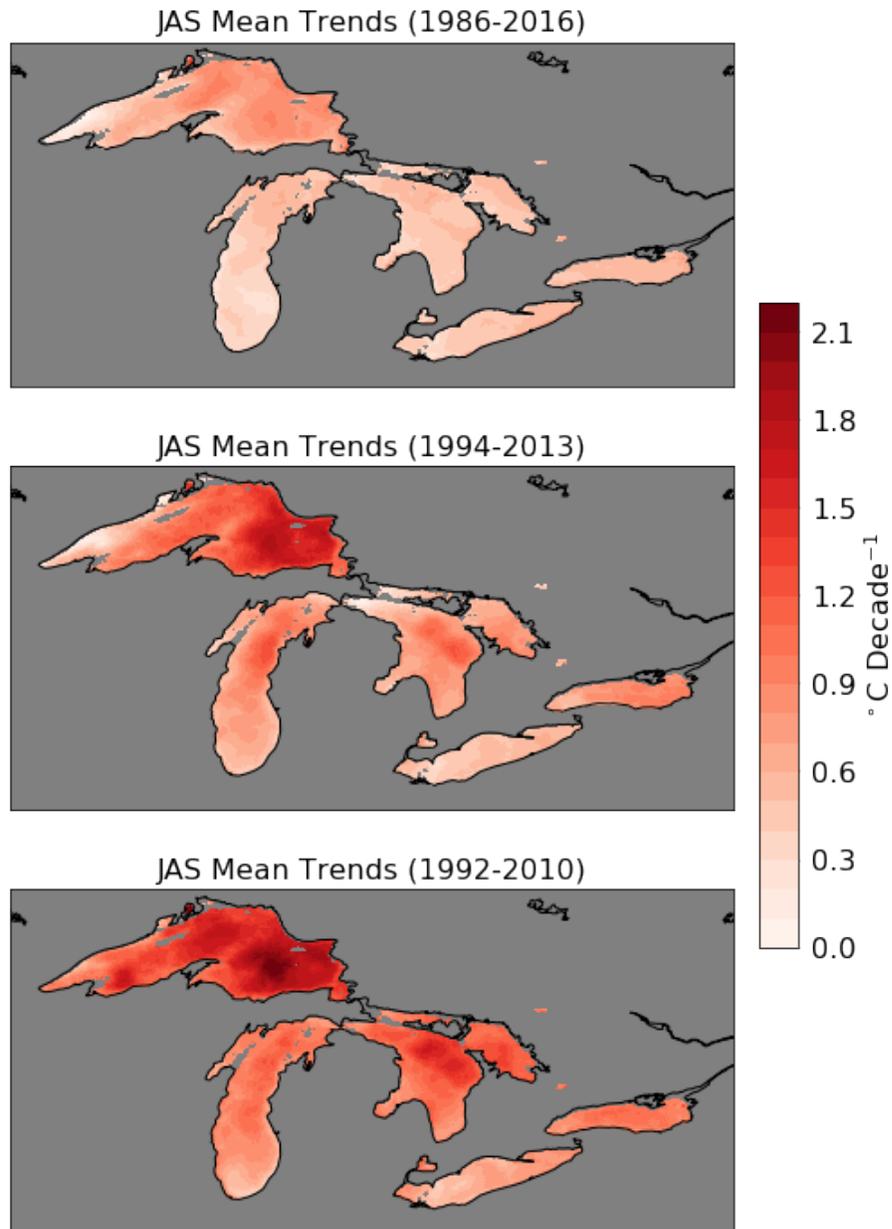


Figure 35. JAS mean trend calculated for three periods (a) the 1986-2016 AVHRR dataset used in this work (b) the 1994-2013 window used in Mason et al. 2016 and (c) the 1992-2010 window used in Woolway et al. 2018.

As presented in Austin and Colman (2007), the distribution of warming rates in Lake Superior appear to be strongly correlated with the bathymetry of the lake. It was further identified in Woolway and Merchant (2018) that surface temperature anomalies persist longer

in the deeper areas of large lakes perhaps contributing to the larger summertime trends seen away from the shoreline in Lake Superior. We have illustrated moderate variability in the correlation of late spring and summertime warming trends and bathymetry between different months of the year. Due to the differing dates of stratification between lakes, the enhanced warming associated with earlier stratification occurs in different months for each lake. For example, there is enhanced warming in the deepest area of lake Michigan in June (Figure 22) relative to shallower areas in June, and the deepest areas in neighboring months. The connection being that the period of maximum intra-annual warming for the deepest areas of Lake Michigan also occurs in mid-to-late June. This same logic follows for Lake Superior. The July warming rates for Superior are largest in the deepest areas (Figure 25 and Figure A.3) and larger than neighboring months. The mean intra-annual warming rates are maximized in mid-to-late July for Lake Superior, and when stratification occurs earlier in the year, the greatest interannual warming are seen in July. To emphasize this point, we expand upon a portion of the analysis in Woolway and Merchant (2018). When comparing the monthly trends between the shallowest 10% and the deepest 10% of for each lake, a similar pattern is observed. For Lake Michigan and Lake Huron, the largest difference in trends during the warming phase of these lakes occurs in June, where the mean intra-annual maximum in warming rate occurs (Figure 16). In August, the differences in trends are small when the maximum temperature is reached (Figure 15), but the overall trend is still positive (Figure 23). The largest difference between the deep and shallow areas of Lake Superior occur in July, again, during the period of maximum intra-annual warming (Figure 16).

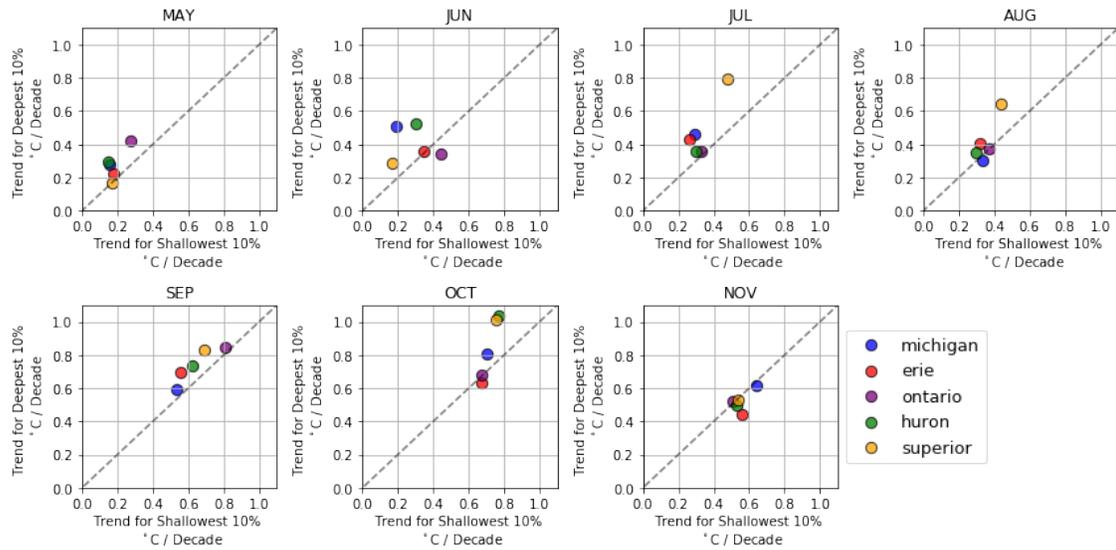


Figure 36. Monthly mean trends (1986-2016) in lake surface temperature shown for each lake for the shallowest 10% and the deepest 10% of the five lakes.

It can be identified in Figure 16 and Figure 17, that the July-September window represents a conceptually different portion of the mean seasonal cycle with respect to intra-annual rates of change for the different lakes in this region. This is particularly the case when comparing the timing of the warming phases of the deepest areas in Lake Michigan and Huron to Lake Superior (Figure 17). A JAS mean would not account for the largely bathymetrically-associated interannual warming occurring in June for Lake Michigan and Lake Huron. Thus, comparing interannual JAS warming rates from Michigan and Huron to Lake Superior, a lake in which the effects of earlier spring stratification are potentially observed in the JAS window, could be misleading in some respects.

The distribution of monthly warming rates shows much more rapid warming trends in specifically September, October and November. This is in agreement with the conclusions of Winslow et al. (2017), which show a similar pattern for much smaller lakes in the same region from 1981 to 2015, but receives relatively little attention in other literature compared to JAS.

In Huron, Superior, and Lake Michigan, there is a second local maximum in correlation between monthly warming rates and bathymetry occurring in October. This is confirmed by examining the difference in October trends between the deep and shallow areas of Superior and Huron. This pattern is relatively muted for Lake Michigan.

The large warming rates that occur during the late summer and early autumn could arise from one of the same mechanisms resulting in enhanced warming in the late spring and early summer. Thermal anomalies tend to persist longer in the deeper areas of the lakes. The period of most rapid cooling is between late September and early October for all five lakes. Given the warming that occurs in the earlier portion of the year we might expect the positive thermal anomalies (with respect to climatology) to persist longer into this mean intra-annual cooling period in the deepest areas. The autumn months also demonstrate lower interannual variability (Figure 18) which could contribute to the differences in long-term trends between JAS and SON given the observed sensitivity to the length, start-year, and end-year of the record (Figure 36). Nonetheless, it is intriguing that the largest trends in lake surface temperatures are observed in the portion of the year with relatively low interannual variability compared to JAS.

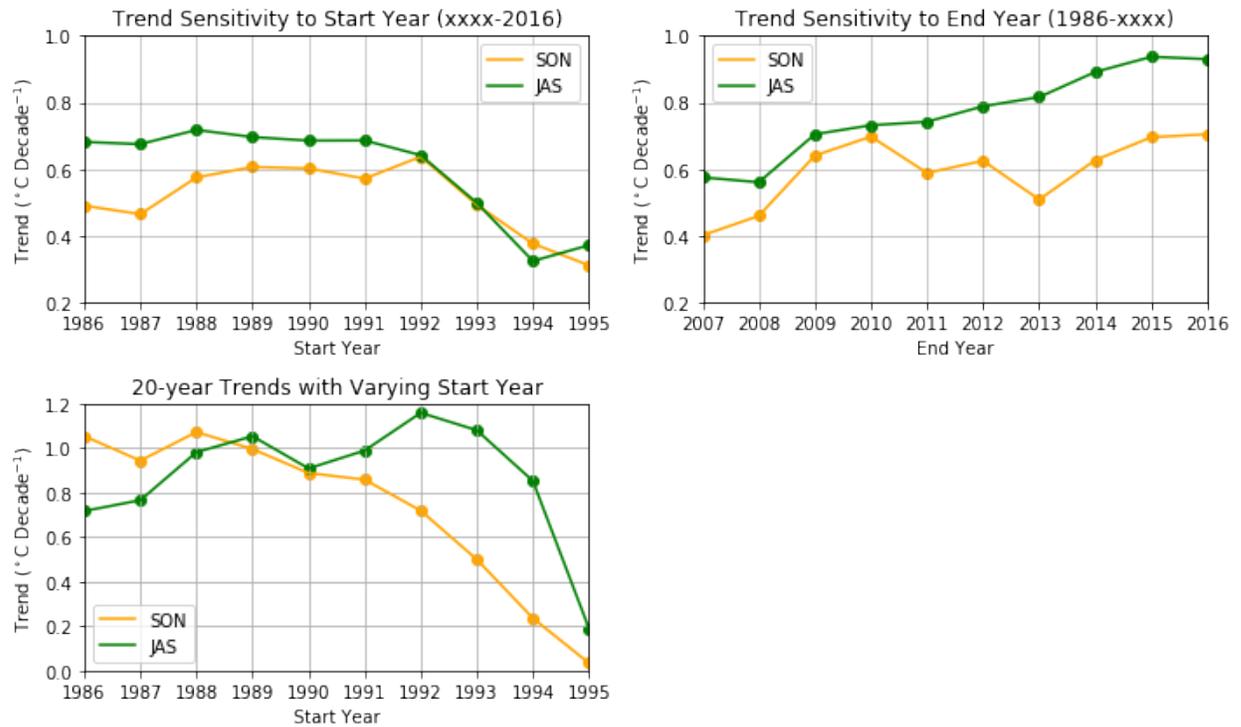


Figure 37. July-September (JAS) and September-November (SON) mean trends in lake surface temperature across the entire Laurentian Great Lakes region calculated for (a) varying start years, (b) varying end years, and (c) a rolling 20-year time series shown at the start year.

To further investigate this sensitivity to the time period examined we calculate both JAS and SON trends for varying start years, end years, and 20-year window for the entire region. Results show a wide range of calculated trends in all of the cases examined. Particularly striking is the JAS trend for 1995-2014, and the JAS trend for 1994-2013, in which starting a time series 1-year earlier results in a 4.5 times increase in the calculated trend ($0.853\text{ }^{\circ}\text{C per decade}$ from $0.19\text{ }^{\circ}\text{C per decade}$). Despite the low interannual variability in lake-wide mean autumn surface temperatures, the SON trends show similar amounts of variability to changing the time window, start year, and end year. However, it is promising the SON trend is less variable when using a longer time series in Figures 37a and 37b.

There are some potential confounding factors in our analysis of long-term trends. The typical temporal separation between clear-sky observations of the HRPT record decreases with the number of operational sensors. This could perhaps result in better capturing nearshore variability later on in the record. At this relatively fine resolution it is plausible that better resolved coastal upwelling events could introduce a negative bias to the trends. This could also affect the spatial patterns seen in warming trends. If small upwelling events are observed more frequently in the latter portion of the record, this could lower the warming trends near the shoreline. However, given the mean duration and frequency of detected events throughout the record, this is unlikely to have a large impact. Most shorelines indicate less than 1 upwelling event per year, and a mean duration of less than 8 days. The largest exception to this is the western shoreline of Lake Michigan where a large number of upwelling events were detected with a mean duration exceeding 10 days.

Table 7. Details of the five datasets from which long term trends are found in Figure 38

Scenario	Diurnal Correction (DC)	Gap Filling (GF)	Observations Included	
			Day	Night
GF-DC-all	X	X	X	X
DC-all	X		X	X
No DC-all			X	X
No DC-day			X	
No DC-night				X

In an attempt to address these potentially confounding factors related to diurnal variability and interpolation, we have calculated long-term trends in Lake Michigan under five scenarios. The five scenarios are described in Table 7. The distribution of warming rates across Lake Michigan on a monthly basis are plotted for each of the scenarios in Figure 38. The general monthly pattern of more rapid warming in the September-November portion of the year still holds across all scenarios, albeit there is small disagreement in magnitude. The largest

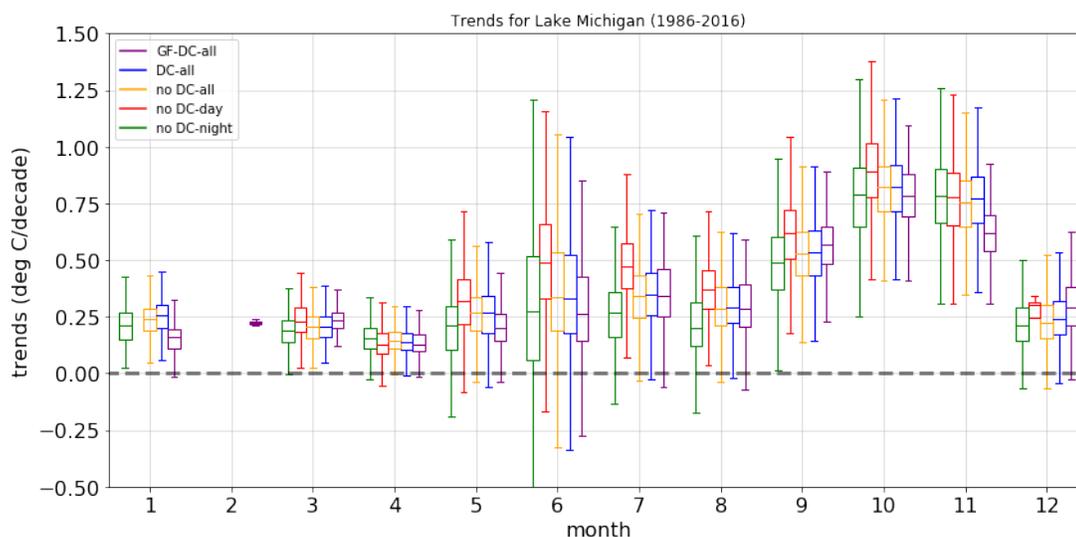


Figure 38. Long-term trends for Lake Michigan using the five different methods detailed in Table 4. Box edges indicate quartiles, the middle line indicates the median, and the whiskers indicate the last values within one interquartile range of the 25th and 75th percentiles.

differences in trends occurs between the use of exclusively daytime and exclusively nighttime imagery. These differences are largest during the summer months and could be a factor of the number of operational AVHRR sensors available in the latter portion of the record that are, as a result, better capturing the peak of the diurnal cycle. Another reason for these differences could be the uneven sampling of the temperatures within months due to the amount observations being effectively halved. The distribution of trends across space appears to be slightly narrower for the gap-filled, diurnally corrected product. This is to be expected since the weighted averages and LOESS will tend to smooth outliers and noise from the time series, and also dampen real observed variability. Overall, there appears to be only small differences between the gap-filled, diurnally-corrected scenario and the scenario with no gap-filling or diurnal correction using the same set of observations. This is unsurprising given our favorable comparison to long-term trends computed in other studies (Mason et al. 2016; Woolway and Merchant 2018).

Spatiotemporal variation was diagnosed by finding EOFs in the deseasonalized and detrended time series. These results show that most of the variation across space is uniform as indicated by the first mode. The implication is that if one region of a lake is warmer or colder than the climatological average, it is more likely than not that the rest of the lake shows a similar pattern. In other words, when the EOF1 index is high, this does not indicate that the entire lake surface is the same temperature, rather, it indicates the lake surface shows somewhat similar departures from climatology. The second mode in Lake Huron and the first mode in Lake Michigan show interesting patterns that appear to be related to the bathymetry of these lakes. This could be representative of the relative response times of the surface temperature anomalies to the forcing acting on them. Woolway and Merchant (2018), as previously mentioned, identified that anomalies tend to persist longer in the deeper areas of the lake. These EOF modes could be a manifestation of this during periods of abrupt warming or cooling, where the deeper areas of the lake tend to respond more slowly than the shallower regions.

Other modes, such as EOF2 in Lake Michigan show strong indication that they are related to coastal upwelling. Examination of the surface temperature anomalies for large positive and large negative values of the EOF 2 index for lake Michigan (not shown) reveals cold anomalies confined to coastal areas on the western and eastern shorelines respectively. This suggests that upwelling could play a large role in enhancing spatiotemporal variation in Lake Michigan.

Motivated by these findings, we developed an upwelling identification scheme adapted from Plattner et al. (2006). Here, we have added a spatial dependence on the region selected for the local temperature distribution used in the Contrast test, rather than predefining the regions as is done in Plattner et al. (2006). We have also added the Cooling and Coastal criteria

for two reasons: first, so we can decrease the threshold for the contrast test— aiming to make this algorithm more sensitive to upwelling, and second, to filter out false positives that are common during the spring onset of stratification along the shoreline.

These criteria suggest that there are several locations on these lakes that experience a high frequency of the large, long-lasting upwelling episodes that are identifiable from satellite imagery. It should be noted that this algorithm does not directly identify the presence of upward vertical motion in the water column. Instead, this algorithm aims to identify regions where upwelling has resulted in abrupt cooling of surface temperatures along the coastline. The distinction between the two being that relatively cold surface temperatures can remain after the upward vertical motion has ceased. It is likely that many upwelling events go undetected by this algorithm for a number of reasons. Most upwelling events have very small spatial extent lasting around 1-2 days (Troy et al. 2012). Overall, this algorithm will only identify extraordinarily large upwelling events resulting in large abrupt surface temperature depressions along the shoreline.

While this method leaves many events undetected, we can be fairly confident that those that are identified are indeed areas where upwelling as occurred, and are forced by the surface wind stress. A cursory analysis of the daily mean wind where increases in upwelling occur, agree with the standard argument of mean Ekman transport to the right of the mean wind stress in large northern hemisphere mid-latitude water bodies. The most convincing portion of this analysis occurs in Lake Michigan. There, the daily mean winds during increases upwelling are almost exclusively parallel to the shoreline with the water to the right of the mean wind direction. These results are in good agreement with a similar, more thorough validation performed for Lake Michigan in Plattner et al. (2006).

We caution against any trend analysis in upwelling, duration, or spatial extent using this dataset. The issue of changing temporal resolution results in a spurious large positive trend in upwelling frequency, and duration. Due to the heavy reliance on interpolation during the earlier portion of this record, it is probable that some upwelling events went unobserved or were smoothed over by the LOESS fit. Therefore this dataset is unlikely to give reliable results on a trend analysis related to the identified upwelling events. Here, we have relied on analysis that examines spatial variability of upwelling frequency and duration that will be less affected by this issue.

8 Conclusions

A long-term (31 years), fine-resolution (0.018°) dataset of daily lake surface water temperature (LSWT) was developed from the Advanced Very High Resolution Radiometer (AVHRR) record for the Laurentian Great Lakes region. In this work, we have made three significant adjustments to the AVHRR LSWT record. Errors in the AVHRR navigation have been identified to be particularly large in the earliest sensors, amounting to a mean bias of over two 0.018° grid cells in the meridional direction in some cases. The removal of these navigation errors results in more accurate LSWT measurements near shorelines. A mean diurnal cycle was found, and subtracted from the LSWT time series to mitigate bias associated with including measurements made at different times of day. Missing LSWT values were filled using a combination of locally weighted interpolation schemes. We have additionally characterized the LSWT observations to have a root mean squared error of 0.588 K compared to buoy measurements. Both the original daily LSWT composites and the gap-filled product will be made available to support fine-scale physical and environmental research in this region. These data can be found through the University of Wisconsin-Madison's Space Science and Engineering Center at <ftp://ftp.ssec.wisc.edu/pub/cwhite/>.

Analysis of the record demonstrates a large amount of variability in warming rates both within and between lakes in this region. When long-term trends are examined on a monthly basis, the largest warming rates are observed to occur in September, October and November, across all five lakes. This has been identified previously in Winslow et al (2017), but shown across space for the largest lakes in this region as is done here. Our results suggest that a comparison of JAS mean trends between different lakes may be somewhat misleading, because

they may represent conceptually different stages of the mean intra-annual variability of surface temperature. Furthermore, the JAS time window does not capture the largest warming trends that are observed in other months. and in some instances there is moderate correlation between the temperature trends and bathymetry, or climatological mean monthly temperature. In all but the shallowest areas of some lakes, we have observed earlier spring stratification. However, a trend sensitivity analysis indicates a large degree of variability in long term trends of surface water temperature.

We have additionally examined spatiotemporal variability in the Laurentian Great Lakes, and found that a large amount of variance in the surface temperature dataset is manifested as spatially uniform departures from climatology. The modes that explain smaller amounts of variability appear to be related to bathymetry, and coastal upwelling. Finally, we have adapted a coastal upwelling identification algorithm, and applied it to the five largest lakes. Here, we have documented the frequency and duration of exceptionally large upwelling events in the Laurentian Great lakes from 1986 to 2016.

9 Bibliography

- Ackerman, S. A., A. Heidinger, M. J. Foster, and B. Maddux, 2013: Satellite Regional Cloud Climatology over the Great Lakes. *Remote Sens.*, **5**, 6223–6240, doi:10.3390/rs5126223.
- Adrian, R., and Coauthors, 2009: Lakes as sentinels of climate change. *Limnol. Oceanogr.*, **54**, 2283–2297, doi:10.4319/lo.2009.54.6_part_2.2283.
<http://www.ncbi.nlm.nih.gov/pubmed/20396409> (Accessed March 12, 2018).
- Anding, D., and R. Kauth, 1970: Estimation of sea surface temperature from space. *Remote Sens. Environ.*, **1**, 217–220, doi:10.1016/S0034-4257(70)80002-5.
<http://linkinghub.elsevier.com/retrieve/pii/S0034425770800025> (Accessed March 13, 2018).
- Assel, R.A. Great Lakes Ice Cover Climatology Update: Winters 2003, 2004, and 2005; NOAA Technical Memorandum GLERL-135; NOAA Great Lakes Environmental Research Laboratory: Ann Arbor, MI, USA, 2005.
- Assel, R.A.; Norton, D.C.; Cronk, K.C. Great Lakes Ice Cover, First Ice, Last Ice, and Ice Duration: Winters 1973–2002; NOAA Technical Memorandum GLERL-121; NOAA Great Lakes Environmental Research Laboratory: Ann Arbor, MI, USA, 2002.
- Austin, J. A., and S. M. Colman, 2007: Lake Superior summer water temperatures are increasing more rapidly than regional temperatures: A positive ice-albedo feedback. *Geophys. Res. Lett.*, **34**, L06604, doi:10.1029/2006GL029021.
<http://doi.wiley.com/10.1029/2006GL029021> (Accessed February 26, 2018).
- Baldwin, D. G., and W. J. Emery, 1993: A systematized approach to AVHRR image navigation. *Ann. Glaciol.*, **17**, 414–420, doi:10.3189/S0260305500013185.
- Binding, C. E., T. A. Greenberg, S. B. Watson, S. Rastin, and J. Gould, 2015: Long term water clarity changes in North America’s Great Lakes from multi-sensor satellite observations. *Limnol. Oceanogr.*, **60**, 1976–1995, doi:10.1002/lno.10146.
- Blumberg, A. F., and D. M. Di Toro, 1990: Effects of Climate Warming on Dissolved Oxygen Concentrations in Lake Erie. *Trans. Am. Fish. Soc.*, **119**, 210–223, doi:10.1577/1548-8659(1990)119<0210:EOCWOD>2.3.CO;2.
[http://doi.wiley.com/10.1577/1548-8659\(1990\)119%3C0210:EOCWOD%3E2.3.CO;2](http://doi.wiley.com/10.1577/1548-8659(1990)119%3C0210:EOCWOD%3E2.3.CO;2) (Accessed June 1, 2018).
- Bordes, P., P. Brunel, A. Marsouin, P. Bordes, P. Brunel, and A. Marsouin, 1992: Automatic Adjustment of AVHRR Navigation. *J. Atmos. Ocean. Technol.*, **9**, 15–27, doi:10.1175/1520-0426(1992)009<0015:AAOAN>2.0.CO;2.
- Carpenter, S. R., E. H. Stanley, and M. J. Vander Zanden, 2011: State of the World’s Freshwater Ecosystems: Physical, Chemical, and Biological Changes. *Annu. Rev. Environ. Resour.*, **36**, 75–99, doi:10.1146/annurev-environ-021810-094524.
<http://www.annualreviews.org/doi/10.1146/annurev-environ-021810-094524> (Accessed March 12, 2018).
- Cleveland, W. S., 1979: Robust locally weighted regression and smoothing scatterplots. *J. Am. Stat. Assoc.*, **74**, 829–836, doi:10.1080/01621459.1979.10481038.
- , S. J. Devlin, and E. Grosse, 1988: Regression by local fitting: Methods, properties, and computational algorithms. *J. Econom.*, **37**, 87–114, doi:10.1016/0304-4076(88)90077-2.
<http://www.sciencedirect.com/science/article/pii/0304407688900772> (Accessed June

- 13, 2018).
- Desai, A. R., J. A. Austin, V. Bennington, and G. A. McKinley, 2009: Stronger winds over a large lake in response to weakening air-to-lake temperature gradient. *Nat. Geosci.*, **2**, 855–858, doi:10.1038/ngeo693. <http://www.nature.com/articles/ngeo693> (Accessed May 30, 2018).
- Donlon, C. J., P. J. Minnett, C. Gentemann, T. J. Nightingale, I. J. Barton, B. Ward, and M. J. Murray, 2002: Toward improved validation of satellite sea surface skin temperature measurements for climate research. *J. Clim.*, **15**, 353–369, doi:10.1175/1520-0442(2002)015<0353:TIVOSS>2.0.CO;2.
- Dudhia, A., 1989: Noise characteristics of the AVHRR infrared channels. *Int. J. Remote Sens.*, **10**, 637–644, doi:10.1080/01431168908903904. <http://www.tandfonline.com/doi/abs/10.1080/01431168908903904> (Accessed June 8, 2018).
- Foster, M. J., and A. Heidinger, 2013: PATMOS-x: Results from a diurnally corrected 30-yr satellite cloud climatology. *J. Clim.*, **26**, 414–425, doi:10.1175/JCLI-D-11-00666.1.
- Gentemann, C. L., C. J. Donlon, A. Stuart-Menteth, and F. J. Wentz, 2003: Diurnal signals in satellite sea surface temperature measurements. *Geophys. Res. Lett.*, **30**, 1140, doi:10.1029/2002GL016291. <http://www.agu.org/pubs/crossref/2003/2002GL016291.shtml>.
- Gentemann, C. L., P. J. Minnett, P. L. Borgne, and C. J. Merchant, 2008: Multi-satellite measurements of large diurnal warming events. *Geophys. Res. Lett.*, **35**, L22602, doi:10.1029/2008GL035730.
- Gillooly, J. F., J. H. Brown, G. B. West, V. M. Savage, and E. L. Charnov, 2001: Effects of size and temperature on metabolic rate. *Science (80-.)*, **293**, 2248–2251, doi:10.1126/science.1061967. <http://www.sciencemag.org/cgi/content/full/293/5538/2248> (Accessed March 12, 2018).
- Haffner, G. D., M. L. Yallop, P. D. N. Hebert, and M. Griffiths, 1984: Ecological Significance of Upwelling Events in Lake Ontario. *J. Great Lakes Res.*, **10**, 28–37, doi:10.1016/S0380-1330(84)71804-1. [http://dx.doi.org/10.1016/S0380-1330\(84\)71804-1](http://dx.doi.org/10.1016/S0380-1330(84)71804-1).
- Heidinger, A. K., A. T. Evan, M. J. Foster, and A. Walther, 2012: A naive Bayesian cloud-detection scheme derived from Calipso and applied within PATMOS-x. *J. Appl. Meteorol. Climatol.*, **51**, 1129–1144, doi:10.1175/JAMC-D-11-02.1. <https://journals.ametsoc.org/doi/pdf/10.1175/JAMC-D-11-02.1> (Accessed February 26, 2018).
- Heidinger, A. K., M. J. Foster, A. Walther, and X. (Tom) Zhao, 2014: The Pathfinder Atmospheres --Extended AVHRR Climate Dataset. *Bull. Am. Meteorol. Soc.*, **95**, 909–922, doi:10.1175/BAMS-D-12-00246.1.
- Kilpatrick, K. a., G. P. Podestá, and R. Evans, 2001: Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *J. Geophys. Res.*, **106**, 9179, doi:10.1029/1999JC000065. <http://dx.doi.org/10.1029/1999JC000065> <http://doi.wiley.com/10.1029/1999JC000065>.
- Kraemer, B. M., and Coauthors, 2015: Morphometry and average temperature affect lake

- stratification responses to climate change. *Geophys. Res. Lett.*, **42**, 4981–4988, doi:10.1002/2015GL064097. <http://doi.wiley.com/10.1002/2015GL064097> (Accessed May 30, 2018).
- , and Coauthors, 2017: Global patterns in lake ecosystem responses to warming based on the temperature dependence of metabolism. *Glob. Chang. Biol.*, **23**, 1881–1890, doi:10.1111/gcb.13459. <https://static1.squarespace.com/static/541825b4e4b0128725f88650/t/57cefc1837c581cd5722846f/1473182747775/23+Kraemer+et+al+2016.pdf> (Accessed March 12, 2018).
- Li, X., W. Pichel, E. Maturi, P. Clemente-Col, and J. Sapper, 2001: Deriving the operational nonlinear multichannel sea surface temperature algorithm coefficients for NOAA-15 AVHRR/3. *int. j. Remote Sens.*, **22**, 699–704, doi:10.1080/01431160010013793. https://pdfs.semanticscholar.org/567d/6f06db52365d3b7d0c973bc47ed256c4977c.pdf?_ga=2.82038422.1267536185.1519234397-495404796.1518577983 (Accessed February 21, 2018).
- Llewellyn-Jones, D. T., P. J. Minnett, R. W. Saunders, and A. M. Zavody, 1984: Satellite multichannel infrared measurements of sea surface temperature of the N. E. Atlantic Ocean using AVHRR / 2. *Q. J. R. Meteorol. Soc.*, **7**, 613–631. http://onlinelibrary.wiley.com/store/10.1002/qj.49711046504/asset/49711046504_ftp.pdf?v=1&t=jeuadm8p&s=a56dbe648ae6bf7dae0d8f3cb995e3b9c02a1b9 (Accessed March 16, 2018).
- Magnuson, J. J., and Coauthors, 2000: Historical trends in lake and river ice cover in the northern hemisphere. *Science*, **289**, 1743–1746, doi:10.1126/SCIENCE.289.5485.1743. <http://www.ncbi.nlm.nih.gov/pubmed/10976066> (Accessed June 1, 2018).
- Mason, L. A., C. M. Riseng, A. D. Gronewold, E. S. Rutherford, J. Wang, A. Clites, S. D. P. Smith, and P. B. McIntyre, 2016: Fine-scale spatial variation in ice cover and surface temperature trends across the surface of the Laurentian Great Lakes. *Clim. Change*, 1–13, doi:10.1007/s10584-016-1721-2. <http://link.springer.com/10.1007/s10584-016-1721-2>.
- Maul, G. A., and M. Sidran, 1972: Comment on “Estimation of Sea Surface Temperature from Space”; *Remote Sens. Environ.*, **2**, 165–169. https://ac.els-cdn.com/0034425771900897/1-s2.0-0034425771900897-main.pdf?_tid=b990929a-c84c-47bf-82fb-c3cfad58eef0&acdnat=1521045413_1abc6e517d7863d50a9cd8f002aca171 (Accessed March 14, 2018).
- McClain, E. P., W. G. Pichel, C. C. Walton, Z. Ahmad, and J. Sutton, 1983: Multi-channel improvements to satellite-derived global sea surface temperatures. *Adv. Sp. Res.*, **2**, 43–47, doi:10.1016/0273-1177(82)90120-X. <https://www.sciencedirect.com/science/article/pii/027311778290120X> (Accessed May 17, 2018).
- McClain, E. P., W. G. Pichel, and C. C. Walton, 1985: Comparative performance of AVHRR-based multichannel sea surface temperatures. *J. Geophys. Res.*, **90**, 11587, doi:10.1029/JC090iC06p11587. <http://doi.wiley.com/10.1029/JC090iC06p11587> (Accessed February 21, 2018).
- McMillin, L. M., 1975: Estimation of Sea Surface Temperatures From Two Infrared Window Measurements With Different Absorption. *J. Geophys. Res.* **DECEMBER**, **80**,

- doi:10.1029/JC080i036p05113.
<http://onlinelibrary.wiley.com/store/10.1029/JC080i036p05113/asset/jgr15200.pdf?v=1&t=jesstccb&s=6eda5dc93e3528ab65882f1f591672aa1d5681c0> (Accessed March 15, 2018).
- McMillin, L. M., and D. S. Crosby, 1984: Theory and Validation of the Multiple Window Sea Surface Temperature Technique. *J. Geophys. Res.*, **89**, 3655–3661, doi:10.1029/JC089iC03p03655. <http://doi.wiley.com/10.1029/JC089iC03p03655> (Accessed March 12, 2018).
- Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Am. Meteorol. Soc.*, **87**, 343–360, doi:10.1175/BAMS-87-3-343. <http://journals.ametsoc.org/doi/10.1175/BAMS-87-3-343> (Accessed May 21, 2018).
- Minnett, P. J., M. Smith, and B. Ward, 2011: Measurements of the oceanic thermal skin effect. *Deep. Res. Part II Top. Stud. Oceanogr.*, **58**, 861–868, doi:10.1016/j.dsr2.2010.10.024. <https://www.sciencedirect.com/science/article/pii/S0967064510003024> (Accessed March 13, 2018).
- Moreno, J. F., and J. Melia, 1993: A Method for Accurate Geometric Correction of NOAA AVHRR HRPT Data. *IEEE Trans. Geosci. Remote Sens.*, **31**, 204–226, doi:10.1109/36.210461.
- North, G. R., T. L. Bell, R. F. Cahalan, F. J. Moeng, G. R. North, T. L. Bell, R. F. Cahalan, and F. J. Moeng, 1982: Sampling Errors in the Estimation of Empirical Orthogonal Functions. *Mon. Weather Rev.*, **110**, 699–706, doi:10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2. <http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281982%29110%3C0699%3ASEITEO%3E2.0.CO%3B2> (Accessed May 17, 2018).
- O'Reilly, C. M., and Coauthors, 2015: Rapid and highly variable warming of lake surface waters around the globe. *Geophys. Res. Lett.*, **42**, 10773–10781, doi:10.1002/2015GL066235. <http://doi.wiley.com/10.1002/2015GL066235> (Accessed August 12, 2016).
- Pareeth, S., N. Salmaso, R. Adrian, and M. Neteler, 2016: Homogenised daily lake surface water temperature data generated from multiple satellite sensors: A long-term case study of a large sub-Alpine lake. *Sci. Rep.*, **6**, 31251, doi:10.1038/srep31251. <http://www.nature.com/articles/srep31251>.
- Pilcher, D. J., G. A. McKinley, H. A. Bootsma, and V. Bennington, 2015: Physical and biogeochemical mechanisms of internal carbon cycling in Lake Michigan. *J. Geophys. Res. Ocean.*, **120**, 2112–2128, doi:10.1002/2014JC010594. <http://doi.wiley.com/10.1002/2014JC010594> (Accessed June 12, 2018).
- Plattner, S., D. M. Mason, G. A. Leshkevich, D. J. Schwab, and E. S. Rutherford, 2006: Classifying and Forecasting Coastal Upwellings in Lake Michigan Using Satellite Derived Temperature Images and Buoy Data. *J. Great Lakes Res.*, **32**, 63–76, doi:10.3394/0380-1330(2006)32[63:CAFUI]2.0.CO;2. [http://www.bioone.org/doi/abs/10.3394/0380-1330\(2006\)32\[63:CAFUI\]2.0.CO;2](http://www.bioone.org/doi/abs/10.3394/0380-1330(2006)32[63:CAFUI]2.0.CO;2).
- Prabhakara, C., V. G. Kunde, and G. Dalu, 1974: Estimation of Sea Surface Temperature From Remote Sensing in the 11- to 13- μ m Window Region. *J. Geophys. Res.*, **79**, 5039–

- 5044, doi:10.1029/JC079i033p05039.
<http://onlinelibrary.wiley.com/store/10.1029/JC079i033p05039/asset/jgr14993.pdf?v=1&t=jercm0q7&s=3f9605b13ea559d5de23f2f8c41e3335dbf802f5> (Accessed March 14, 2018).
- Rahel, F. J., and J. D. Olden, 2008: Assessing the effects of climate change on aquatic invasive species. *Conserv. Biol.*, **22**, 521–533, doi:10.1111/j.1523-1739.2008.00950.x.
- Rao, Y. R., and C. R. Murthy, 2001: Coastal Boundary Layer Characteristics during Summer Stratification in Lake Ontario. *J. Phys. Oceanogr.*, **31**, 1088–1104, doi:10.1175/1520-0485(2001)031<1088:CBLCDS>2.0.CO;2.
<https://journals.ametsoc.org/doi/pdf/10.1175/1520-0485%282001%29031%3C1088%3ACBLCDS%3E2.0.CO%3B2> (Accessed February 20, 2018).
- Reynolds, R. W., 1993: Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures. *J. Clim.*, doi:10.1175/1520-0442(1993)006<0768:IOMPAO>2.0.CO;2.
- Riffler, M., G. Lieberherr, and S. Wunderle, 2015: Lake surface water temperatures of European Alpine lakes (1989–2013) based on the Advanced Very High Resolution Radiometer (AVHRR) 1 km data set. *Earth Syst. Sci. Data*, **7**, 1–17, doi:10.5194/essd-7-1-2015. www.earth-syst-sci-data.net/7/1/2015/ (Accessed August 12, 2016).
- Schindler, D. W., 2009: Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscapes. *Limnol. Oceanogr.*, **54**, 2349–2358, doi:10.4319/lo.2009.54.6_part_2.2349.
http://doi.wiley.com/10.4319/lo.2009.54.6_part_2.2349 (Accessed June 1, 2018).
- Schluessel, P., W. J. Emery, H. Grassl, and T. Mammen, 1990: On the bulk-skin temperature difference and its impact on satellite remote sensing of sea surface temperature. *J. Geophys. Res.*, **95**, 13341, doi:10.1029/JC095iC08p13341.
<http://doi.wiley.com/10.1029/JC095iC08p13341> (Accessed April 9, 2018).
- Schneider, P., and S. J. Hook, 2010: Space observations of inland water bodies show rapid surface warming since 1985. *Geophys. Res. Lett.*, **37**, doi:10.1029/2010GL045059.
- Schwab, D. J., G. A. Leshkevich, and G. C. Muhr, 1999: Automated mapping of surface water temperature in the Great Lakes. *J. Great Lakes Res.*, **25**, 468–481, doi:10.1016/S0380-1330(99)70755-0.
<https://www.sciencedirect.com/science/article/pii/S0380133099707550> (Accessed February 23, 2018).
- Smith, A. L., and Coauthors, 2012: Effects of climate change on the distribution of invasive alien species in Canada: a knowledge synthesis of range change projections in a warming world. *Environ. Rev.*, **20**, 1–16, doi:10.1139/A11-020.
<http://www.nrcresearchpress.com/doi/pdf/10.1139/a11-020> (Accessed June 1, 2018).
- Strong, A. E., and E. P. McClain, 1984: Improved ocean surface temperature from space: comparison with drifting buoys. *Bull. rhe Am. Meteorol. Soc.*, **85**, 138–142.
- Troy, C. D., S. Ahmed, N. Hawley, and A. Goodwell, 2012: Cross-shelf thermal variability in southern Lake Michigan during the stratified periods. *J. Geophys. Res. Ocean.*, **117**, 1–16, doi:10.1029/2011JC007148.
- Walton, C. C., 1988: Nonlinear Multichannel Algorithms for Estimating Sea Surface Temperature with AVHRR Satellite Data. *J. Appl. Meteorol.*, **27**, 115–124, doi:doi:10.1175/1520-0450(1988)027<0115:NMAFES>2.0.CO;2.

- <http://journals.ametsoc.org/doi/abs/10.1175/1520-0450%281988%29027%3C0115%3ANMAFES%3E2.0.CO%3B2>.
- Walton, C. C., W. G. Pichel, J. F. Sapper, and D. A. May, 1998: The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with the NOAA polar-orbiting environmental satellites. *J. Geophys. Res. Ocean.*, **103**, 27999–28012, doi:10.1029/98JC02370.
<http://doi.wiley.com/10.1029/98JC02370> (Accessed March 13, 2018).
- Wang, J.; Assel, R.A.; Walterscheid, S.; Clites, A.H.; Bai, X. Great Lakes Ice Climatology Update, Winters 2006–2011, Description of the Digital Ice Cover Dataset; NOAA Technical Memorandum GLERL-155; NOAA Great Lakes Environmental Research Laboratory: Ann Arbor, MI, USA, 2012.
- Warren, D., 1989: AVHRR channel-3 noise and methods for its removal. *Int. J. Remote Sens.*, **10**, 645–651, doi:10.1080/01431168908903905.
<http://www.tandfonline.com/doi/abs/10.1080/01431168908903905> (Accessed June 8, 2018).
- Webster, P. J., C. A. Clayson, J. A. Curry, P. J. Webster, C. A. Clayson, and J. A. Curry, 1996: Clouds, Radiation, and the Diurnal Cycle of Sea Surface Temperature in the Tropical Western Pacific. *J. Clim.*, **9**, 1712–1730, doi:10.1175/1520-0442(1996)009<1712:CRATDC>2.0.CO;2.
<http://journals.ametsoc.org/doi/abs/10.1175/1520-0442%281996%29009%3C1712%3ACRATDC%3E2.0.CO%3B2> (Accessed April 9, 2018).
- Wilson, R. C., S. J. Hook, P. Schneider, and S. G. Schladow, 2013: Skin and bulk temperature difference at Lake Tahoe: A case study on lake skin effect. *J. Geophys. Res. Atmos.*, **118**, 10332–10346, doi:10.1002/jgrd.50786.
- Winslow, L. A., J. S. Read, G. J. A. Hansen, K. C. Rose, and D. M. Robertson, 2017: Seasonality of change: Summer warming rates do not fully represent effects of climate change on lake temperatures. *Limnol. Oceanogr.*, **62**, 2168–2178, doi:10.1002/lno.10557. <http://doi.wiley.com/10.1002/lno.10557> (Accessed March 13, 2018).
- Woolway, R. I., and C. J. Merchant, 2018: Intralake Heterogeneity of Thermal Responses to Climate Change: A Study of Large Northern Hemisphere Lakes. *J. Geophys. Res. Atmos.*, **123**, 3087–3098, doi:10.1002/2017JD027661.
- , and Coauthors, 2016: Diel surface temperature range scales with lake size. *PLoS One*, **11**, 1–14, doi:10.1371/journal.pone.0152466.
- Xu, F., and A. Ignatov, 2014: In situ SST quality monitor (iQuam). *J. Atmos. Ocean. Technol.*, **31**, 164–180, doi:10.1175/JTECH-D-13-00121.1.
<https://journals.ametsoc.org/doi/pdf/10.1175/JTECH-D-13-00121.1> (Accessed March 12, 2018).
- Zhong, Y., M. Notaro, S. J. Vavrus, and M. J. Foster, 2016: Recent accelerated warming of the Laurentian Great Lakes: Physical drivers. *Limnol. Oceanogr.*, **61**, 1762–1786, doi:10.1002/lno.10331. <http://doi.wiley.com/10.1002/lno.10331> (Accessed July 22, 2016).

10 Appendix

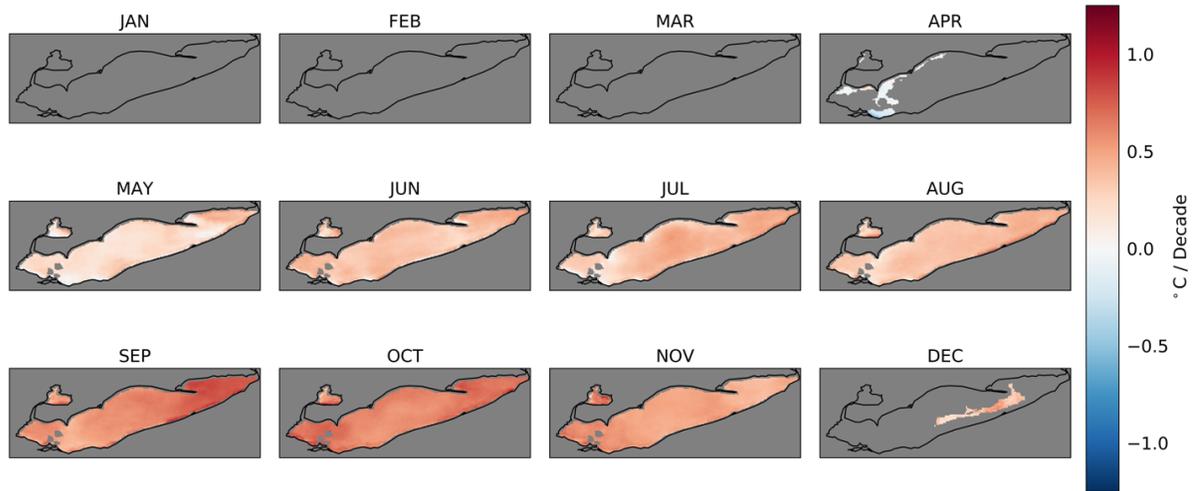


Figure A.2. Trends in monthly mean temperature for Lake Erie. Trends are only shown for months and locations where more than 28 years are ice-free.

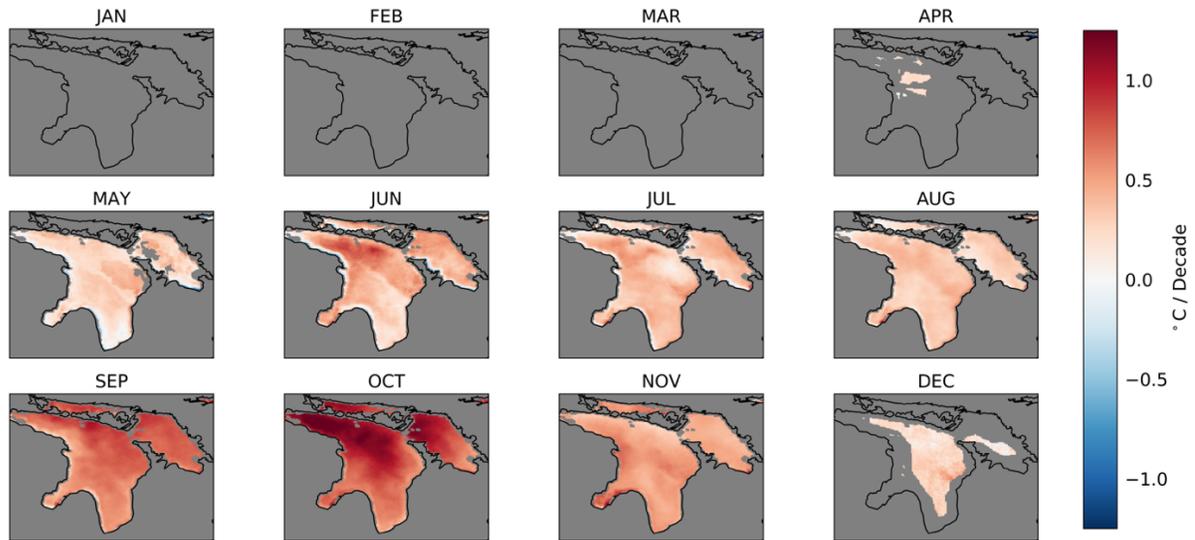


Figure A.2. Trends in monthly mean temperature for Lake Huron. Trends are only shown for months and locations where more than 28 years are ice-free.

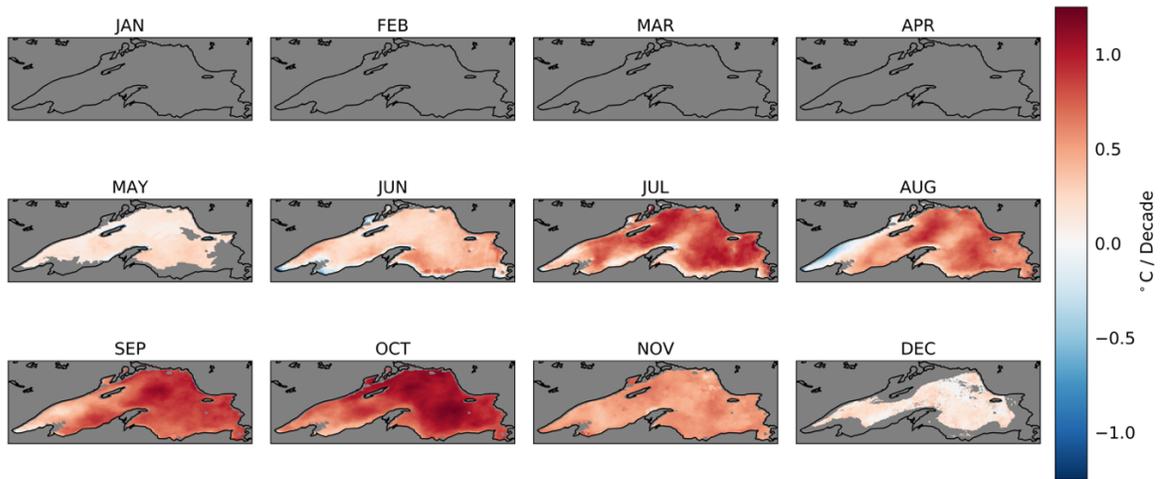


Figure A.3. Trends in monthly mean temperature for Lake Superior. Trends are only shown for months and locations where more than 28 years are ice-free.

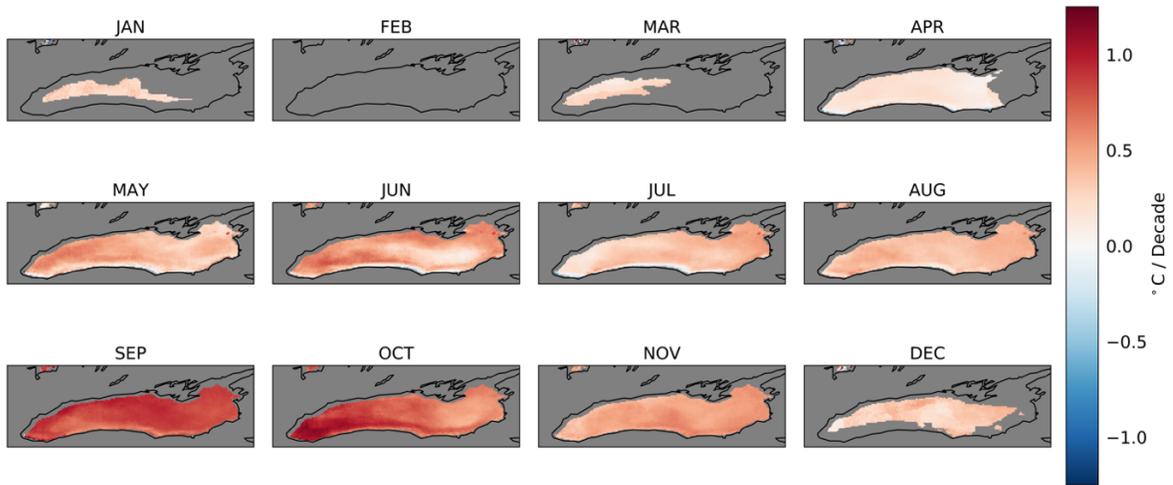


Figure A.4 Trends in monthly mean temperature for Lake Ontario. Trends are only shown for months and locations where more than 28 years are ice-free.

