Brigham Young University
BYU ScholarsArchive

Theses and Dissertations

2019-07-01

Landsat Collections Reveal Long-Term Algal Bloom Hot Spots of Utah Lake

Rachel Shanae Tate
Brigham Young University

Follow this and additional works at: https://scholarsarchive.byu.edu/etd

BYU ScholarsArchive Citation
Tate, Rachel Shanae, "Landsat Collections Reveal Long-Term Algal Bloom Hot Spots of Utah Lake" (2019). *Theses and Dissertations*. 8585.
https://scholarsarchive.byu.edu/etd/8585

This Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Landsat Collections Reveal Long-Term Algal Bloom Hot Spots of Utah Lake

Rachel Shanae Tate

A thesis submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Neil C. Hansen, Chair
Mark W. Jackson
Zachary T. Aanderud

Department of Plant and Wildlife Sciences

Brigham Young University

Copyright © 2019 Rachel Shanae Tate

All Rights Reserved

ABSTRACT

Landsat Collections Reveal Long-Term Algal Bloom Hot Spots of Utah Lake

Rachel Shanae Tate
Department of Plant and Wildlife Sciences, BYU
Master of Science

Harmful algal blooms (HABs) and nuisance algal blooms (NABs) are a worldwide phenomenon with implications for human health and safety. HABs occur when algae or bacteria grow in high enough densities to harm animals and humans. A primary component of harmful algal blooms is cyanobacteria, which are aquatic, photosynthesizing microorganisms that produce toxins at high concentrations. Cyanobacterial biomass has increased worldwide in recent decades, raising concern about the future of fresh- and marine-water systems in a changing climate. Understanding the patterns and conditions of past algal blooms can provide useful insights for managing future blooms. Remote sensing can enhance our understanding of the spatiotemporal distribution of HABs and NABs. We used radiometrically corrected images from the USGS Landsat Collections available in the Google Earth Engine for cloud processing. In 2016, the USGS calibrated the sensors of Landsat 4, 5, 7, and 8 to create a continuous collection of satellite images from 1984 to present. We use this 34-year dataset to expand the historical record of algal blooms at our study site and to understand factors relating to the spatiotemporal patterns of these blooms. We applied three models, including the Floating Algae Index (FAI), the Normalized Difference Vegetation Index (NDVI), and one developed with *in situ* chlorophyll-a (chl-a) data, to 398 images masked for cloud cover and lake elevation taken from 34 growing seasons (April – October). We found that the Normalized Difference Water Index (NDWI) used to separate water and land pixels fails under algal bloom conditions, whereas a modified NDWI does not. We also performed an emerging hot spot analysis in ArcGIS using the chlorophyll-a, NDVI, and FAI predictions from the surface reflectance values of the images. Our analysis indicates that the Provo Bay and parts of the eastern shoreline of Utah Lake have had algal blooms for 30 out of the 34 years included in this study, rendering them enduring hot spots. The remainder of the lake is a cold spot, showing clusters of low mean chl-a, NDVI, and FAI values over time. The overall trend of mean NDVI and lake surface area over this 34-year dataset is decreasing, whereas lake water temperature is increasing. This study develops a method for analyzing algal blooms over multiple decades and provides useful information for the management and prediction of future blooms.

Keywords: algae, cyanobacteria, HABs, Landsat, Earth Engine, hot spot analysis

ACKNOWLEDGEMENTS

I would like to thank my family and friends for their consistent support throughout my educational endeavors and my entire life. They have encouraged and helped me to achieve my goals, overcome my fears, and keep pushing forward. I appreciate all that they have done to help me be where I am today.

I also express my gratitude for my committee and for all of the guidance, suggestions, and encouragement they gave me. I could not have done this without them. I am especially grateful for all of the encouragement and advice Dr. Hansen has given me through both my undergraduate and graduate degrees. He has been a great friend and mentor to me and I will always be grateful for that. I would also like to thank Dr. Jackson for his GIS wizardry and for all that he has taught me about GIS, computers, and research. He helped me discover my love for GIS and helped me solve problems that had me banging my head against the wall. I would also like to thank all of the faculty and staff on campus that have provided invaluable assistance that makes this research possible.

Lastly, I extend special thanks to Steve Schill from the Nature Conservancy and Jess Clark and Josh Goldstein from the USDA Forest Service Remote Sensing Application Center for their introduction to Google Earth Engine. I am also grateful to Ben Holcomb at the Utah Department of Water Quality, and Carly Hansen for their help in formulating this project. Each of these people guided me to valuable questions and solutions for this research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Harmful algal blooms (HABs) and nuisance algal blooms (NABs) are a world-wide phenomenon affecting the health of humans and animals. Cyanobacteria are a key component in both HABs and NABs. Cyanobacterial biomass has increased in recent decades on multiple continents, rendering them a growing global phenomenon (Paerl and Paul 2012; Taranu et al. 2015). Cyanobacteria usually produce intracellular cyanotoxins during death or cell rupture and some cyanobacteria release extracellular toxins without cell rupture or death (EPA 2018). Different toxins, produced by different species of cyanobacteria, may cause gastrointestinal, neurological, liver, kidney, and respiratory impairments (Clark et al. 2017). Because HABs have a serious implication for human health, scientists are employing many different methods to monitor and predict their activity. Some of these methods can be used to create a history of algal growth, which can then allow exploration of patterns and relationships between algae and other controlling variables (Ho et al. 2017).

Geographic Information Systems (GIS) and remote sensing tools can be used to create a historical record of temporal and spatial patterns in algal density. GIS integrates multiple factors influencing algal blooms, for example, remote sensing tools are widely used to quantify water quality characteristics, such as secchi disk transparency, colored dissolved organic matter, and chlorophyll (Brezonik et al. 2005). In terms of algal blooms, publically available data from satellites are commonly used in remote sensing applications to model outbreak and intensity (Brezonik et al. 2005; Ho et al. 2017; Stumpf et al. 2016), but historical analyses of algal blooms using remote sensing is not common. Recent advances render an historical analysis of using satellite data more approachable. In 2016 the USGS compiled a tiered collection of satellite images composed of data from Landsat 4-5 Thematic Mapper, Landsat 7 Enhanced Thematic

Mapper Plus, and Landsat 8 Operational Land Imager/Thermal Infrared Sensor instruments (USGS 2017). The Landsat Collections mark a significant change in the ability to process and access Landsat data. The data of the highest quality are suitable for time-series analysis and are placed in the Tier 1 Collection (USGS 2017). The calibration between sensors simplifies the processing time of the images and expands the range of dates possible in historical analyses. Additionally, Google recently developed the Earth Engine that significantly reduces the time required to process images. The Google Earth Engine hosts the entire Landsat Collection, as well as public data from Sentinel, Terra, Aqua, and other satellites. Satellite scenes can be processed within seconds using cloud computing (Gorelick et al. 2017), drastically reducing the image processing time by eliminating the need to download and radiometrically correct each scene individually. In the context of this study, data from these satellites can be used with spatial statistics to find locations with significant algal densities, or "hot spots" over time (Esri 2018c; Harris et al. 2017). This study demonstrates how Landsat Collection 1, Google Earth Engine, and Emerging Hot Spot Analysis can work together as a unique combination of tools to generate a historical analysis of algal blooms and a better understanding of the factors that influence them.

Multiple natural and anthropogenic variables influence algal growth. Some abiotic factors are climate and water level (da Costa et al. 2016; Urquhart et al. 2017). Notable climate variables include changes in extreme weather events and increased water temperature (Urquhart et al. 2017). Other main variables include water residence time and light intensity, which is governed by solar radiation (Fadel et al. 2015). The season in which most HABs in North America occur is between June and October, when air and water temperatures are warm and solar radiation is high (Bertani et al. 2017). Reductions in water level can increase nutrient concentration and alter light availability (da Costa et al. 2016). Changes in nutrient availability can favor cyanobacteria

that can outcompete other phytoplankton using the same resources. Combinations of these factors can influence the growth of cyanobacteria in aquatic systems.

Utah Lake, located east of the Wasatch Range, had HAB warning advisories during the summers of 2016, 2017, and 2018 (DEQ 2016, 2017, 2018). Multiple beaches around the lake closed as a result of the advisories, discouraging people from participating in recreational activities on the lake. Several cases of animal ailments, including dog mortality have been linked to HABs on Utah Lake (DEQ 2014, 2016). Although Utah Lake's recent toxic cyanobacteria blooms are causing alarm, there are records that suggest that HABs are not new, including one publication detailing a cyanobacterial bloom as early as 1972 (Strong 1974). There is interest to determine whether algal blooms in Utah Lake occur more or less frequently than they have in the past, but available water quality data are inadequate to answer this question. A record of past blooms obtained with remote sensing can be employed to enhance the information gleaned from the increased monitoring that has accompanied the recent alarm (Page et al. 2018).

The objective of this study was to develop a method for studying spatiotemporal algal bloom dynamics using the Landsat Archive, Google Earth Engine, and hot spot analysis. We use Utah Lake as a case study for this method of finding hot spots of algae growth over 35 years. Since toxins have no pigments, they cannot be detected directly from satellites, thus this paper addresses the detection of algal blooms in general and not HABs specifically. Further, due to the limited spectral resolution of Landsat satellites and lack of a red edge band, this study does not seek to distinguish between cyanobacteria and green algae. We hypothesized that the frequency and extent of algal blooms at our study site do not increase between 1984 and 2018, but that hot spots for algal growth occur near locations where wastewater effluent discharges into the lake. We also hypothesized that algal blooms are most likely to occur during hot and dry summers

when lake water level is low and lake water temperature is relatively high, as cyanobacteria biomass can increase with temperatures up to 19°C (Walls et al. 2018). The relationship between these factors and the spatial distribution of historical outbreaks will provide valuable insight behind these blooms that can be used to improve the prediction of future blooms and management approaches to limit them.

METHODS

Site Description

Utah Lake is a shallow, hypereutrophic lake in north-central Utah, U.S.A. in the Basin and Range province (PSOMAS 2007). The lake has a maximum depth of 4.2 m, an average secchi depth of 12 cm (Miller and Crowl 2006), and an average surface area of 380 km^2, designating Utah Lake as one of the largest naturally occurring freshwater lakes in the Western U.S.A. (PSOMAS 2007). The Utah Lake watershed receives an average of 47 cm of annual precipitation, with the majority falling in the form of snow in the mountainous areas of the watershed. Surface water from the Uinta and Weber Basins is imported to the Utah Lake watershed to help meet growing urban water demands (DWR 2014).

Utah Lake has a unique chemistry that alters its optical properties, rendering difficulty when performing spectral analyses using satellite imagery. Utah Lake is a turbid lake characterized by mineral deposits throughout the water column. The sediments are dominated by the mineral calcite ($CaCO_3$) and have a high sorption capacity for Phosphorus (Randall et al. 2019). The mixing of these sediments through the water column adds to the turbidity of the lake and the availability of nutrients. The concentrations of total P in lake sediments ranges from 280 ppm to 1710 ppm (Randall et al. 2019). Suspended particles in the water column alter the reflection of

light. Brightness in Landsat's visible and near-infrared wavelengths increase as water clarity decreases due to scattering (Kloiber et al. 2002). Natural and anthropogenic sources discharge P into the lake, including 7 wastewater treatment plants on the east side of the lake (Figure 1).

Image Processing and Acquisition

We processed surface reflectance images from Landsat Collection 1 Tier 1 and prepared them for analysis using the Google Earth Engine. Collection 1 Tier 1 data meet strict radiometric and geometric quality requirements and are the highest-quality Landsat data products available. Surface Reflectance data available in the Google Earth Engine are radiometrically corrected using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) software developed by the USGS (Google 2017). LEDAPS performs a simple cloud shadow mask as well as atmospheric correction using ozone, temperature, water vapor, and a DEM (Schmidt et al. 2013). We filtered the collection of surface reflectance images from Landsat 5, 7, and 8 in the Earth Engine Code Editor to include images of Utah Lake acquired between the months of April and October. Then the images were filtered for clouds and masked for both clouds and water.

Cloud Filters and Masks

Our study required development of a method for filtering cloudy images to reduce error in image analysis. We compared different methods of removing cloudy images, which included using limited scene cloud cover and limited cloud cover over just the Utah Lake study site. The Scene Cloud Cover Score was calculated using the C Function of Mask (CFMask) algorithm, which is an improved version of the FMask function (USGS 2018b). This score calculates the percentage of cloud cover in an entire Landsat scene, which occupies about 35,150 km^2. To filter

images with clouds just over the lake, we masked the cloudy pixels using the Quality Assessment Band (pixel_qa) and counted the remaining pixels inside the Utah Lake compromise line. The compromise line is the lake surface elevation at which the gates that release water from Utah Lake to the Jordan River must be fully open to prevent flooding (DWR 2014). The pixel_qa band is included with Landsat Collection 1 Level-1 products. This band designates pixels as cloud, cloud shadow, snow, ice, or cirrus clouds using the CFMask algorithm. The CFMask algorithm is one of the best available algorithms for identifying and masking clouds from Landsat imagery (Foga et al. 2017; USGS 2018a).

Water Mask Comparison

We evaluated four different water masks to account for a dynamic shoreline due to changes in precipitation and water usage. The four water mask approaches used variations of the Normalized Difference Water Index (NDWI) and the near infrared band (NIR) reflectance values (McFeeters 1996; Xu 2006). The masks were the Normalized Difference Water Index, NIR, NDWI with NIR, and the Modified Normalized Difference Vegetation Index (MNDWI). The Normalized Difference Water Index (NDWI; McFeeters 1996; Equation 1) normally gives values greater than zero for water and values less than zero for land.

$$\mathrm{NDWI} = \frac{\mathrm{Green}-\mathrm{NIR}}{\mathrm{Green}+\mathrm{NIR}} \quad (1)$$

The NIR mask was developed empirically and uses NIR values < 1200 as water because it absorbs near infrared light. The NDWI with NIR designates pixels with either a positive NDWI value or a low NIR value (< 1200) as water. The final mask was the Modified Normalized Difference Water Index (MNDWI; Equation 2), which uses a middle infrared band instead of the near infrared band (Xu 2006).

$$\text{MNDWI} = \frac{\text{Green}-\text{SWIR1}}{\text{Green}+\text{SWIR1}} \quad (2)$$

When applied to scenes from Landsat 5 and 7, the NDWI uses Band 4, whereas the MNDWI uses Band 5. The four masks were evaluated using 11 images known to have some degree of algal bloom present and 11 randomly selected images. We visually compared the result of the water mask with the true color satellite images and quantified the total lake area classified as water. Once we found a suitable water mask, we applied it to each satellite image used in the study. We then used the masked images to predict algal density on the lake using three models.

Model Descriptions

We used the models calibrated to our study site to predict chlorophyll-a (chl-a) concentration as an approximation of algal biomass (Hansen et al. 2019; Noack et al. 1985; Sadeghian et al. 2018; Tebbs et al. 2013). The chl-a models were calibrated using field measurements of Utah Lake from the Ambient Water Quality Monitoring System (AWQMS) database from the Utah Department of Environmental Quality, Division of Water Quality (UT-DWQ). The surface reflectance values from the images were correlated to the chlorophyll-a field data using a generalized linear model structure. Model development was characterized by an empirical approach that considered all of the bands and band ratios for Landsat 5 and 7. Separate models for the early and late summer months account for changes in algal community structure throughout the growing season (Whiting et al. 1978). The models for two unique time periods are expressed in Equations 3 and 4 below:

May-June Chl = exp(-14.23 + 9.33*Green/Blue + 0.003*Blue - 0.004*SWIR1) (3)

July-Sept Chl = exp(7.33 - 0.004*Blue - 0.05*Green/SWIR2 + 0.01*Red/SWIR1), (4)

where SWIR1 and 2 represent the shortwave infrared bands.

In addition to the models predicting chlorophyll-a concentration, we used the Normalized Difference Vegetation Index (NDVI) and the Floating Algae Index (FAI) as approximations for algal density. The NDVI is commonly used for algal detection due to the versatility and simplicity of using a normalization between the red and near infrared bands. The simplicity of the NDVI enables a continuous spatiotemporal analysis between the Landsat TM and MSS (Landsat 1, 2, and 3) satellites (Augusto-Silva et al. 2014; Ogashawara et al. 2016; Oyama et al. 2015). The Floating Algae Index is also widely used to estimate the density of algae floating on the water (Hu 2009; Oyama et al. 2015; Stumpf et al. 2016). The FAI model as defined by Hu (2009) is described in Equation 5:

$$FAI = NIR - [Red + (SWIR - Red) \times (\lambda_{NIR} - \lambda_{Red})/(\lambda_{SWIR} - \lambda_{Red})], \quad (5)$$

where λ is the wavelength midpoint for each band. For Landsat 4, 5, and 7, λ_{NIR} = 830 nm, λ_{Red} = 660 nm, and λ_{SWIR} = 1650 nm. For Landsat 8, λ_{NIR} = 865 nm, λ_{Red} = 655 nm, and λ_{SWIR} = 1609 nm.

Model Application

Chlorophyll-a, NDVI, and FAI predictions for Landsat 5-8 were extracted to points located 90 m apart in the Google Earth Engine. The 90 m spacing improved consistency between MSS, TM, and OLI satellites as well as faster processing time outside of Earth Engine. The points were created using a Landsat 8 image which was resampled to 90 m with the "Raster to Point" tool in ArcGIS Desktop. A shapefile of these points was uploaded as an asset to the Earth Engine. The NDVI, chlorophyll-a, and FAI values were extracted to the points using the image.reduceRegions() function. These enriched points were exported for further analysis in ArcGIS Pro 2.3, ArcMap 10.6, and R 3.2.4.

Hot Spot Analysis

We prepared the data to find locations where algal concentrations were consistently high over 34 years in emerging hot spot analysis. The data were prepared for hot spot analysis using the Create Space Time Cube by Aggregating Points tool in ArcGIS Pro. The exported points were projected into the WGS 1984 UTM 12 N projection and were aggregated by year into 180 m square space-time bins. Square bins with sides 180 m long include 4 points in each bin. We also specified a reference time of 1 November 2018 to ensure consistent alignment of bins with the end of the water year instead of the end of the image record for each year. Empty bins were filled with spatial neighbors. This parameter includes neighbors and neighbors of neighbors for both the edges and nodes of the bin in the Queens Case Contiguity method (Figure 2). At least four neighbors are required to fill an empty bin with a spatial neighbor (Esri 2018a). The Mann-Kendall statistic is calculated to measure the temporal trend at each bin (Esri 2018b). The output of the Create Space Time Cube by Aggregating Points tool is a Network Common Data Form (netCDF) file. The netCDF file is then used as the input for the Emerging Hot Spot Analysis tool. The Emerging Hot Spot Analysis tool finds hot and cold spots of algal growth on the lake. A hot spot is a location that exhibits a clustering of high values that is statistically significant. Here, hot spots represent areas that have clusters of high chl-a, NDVI, or FAI values due to underlying spatial processes and not to random chance (Esri 2018c; Harris et al. 2017). The Emerging Hot Spot Analysis tool finds hot spots and compares them over time to produce 17 categories of hot or cold spots. This tool works by calculating the degree of clustering and the temporal trends using the Getis-Ord Gi* and Mann-Kendall statistics, respectively (Harris et al. 2017; Mann 1945; Ord and Getis 1995).

The Getis-Ord Gi* statistic calculates the degree of clustering by calculating local and global sums. The sum of a neighborhood of points with high NDVI values, or local sum, is compared to the sum of all NDVI values, or global sum. If the difference between the local and global sums is too large to be a result of random chance, the point has a statistically significant *z*-score ($-1.96 < z < 1.96$). Statistically significant z-scores signify the presence of a hot or cold spot and translate to significant *p*-values ($p < 0.05$). For $z > 1.96$, the larger the *z*-score, the more intense the hot spot. For $z > 1.96$, the smaller the *z*-score, the more intense the cold spot (Esri 2018c, d). Because the data have spatial and temporal neighbors, the spatial neighborhood size was 361 m to include 2 neighboring bins in each direction and 1, 4, 9, and 19 previous time steps for the temporal neighborhood (Harris et al. 2017).

After the Getis-Ord Gi* statistic is calculated and added to the input netCDF file, the Emerging Hot Spot Analysis tool uses the Mann-Kendall trend test to analyze temporal patterns. The Mann-Kendall trend test is a non-parametric application of Kendall's rank order correlation test (Harris et al. 2017). The test compares each bin with the bin in the previous time step and assigns a value of +1, -1, or 0 if the second is greater, less than, or equal to the first, respectively. These ranks are summed and used to determine if the difference between the sum and zero is significant or not. A bin with a positive z-score indicates an increasing trend for that bin over time (Esri 2018b).

The Emerging Hot Spot Analysis tool separates bins into persistent, emerging, sporadic, oscillating, and historical hot and cold spots, among other categories. We combined several categories to simplify the display and interpretation of results (Table 1). The intensifying category remained unchanged. The categories where less than 90% of bins were statistically significant were combined into a new category that we named "dynamic." The categories that we

renamed to “dynamic” were new, sporadic, oscillating, and consecutive hot and cold spots. Categories where greater than 90% of the bins were statistically significant hot or cold spots were combined and renamed “enduring” hot or cold spots. Enduring hot spots include categories previously known as persistent or historical hot or cold spots.

Correlation Analysis of Hot Spots to Water Level and Temperature to Understand Trends

We acquired measures of water temperature and lake surface area from the Landsat imagery to find relationships between blooms, water level, and temperature. Landsat 5, 7, and 8 products include thermal bands. Lake surface area was used as an approximation of water level and was estimated using images that had been masked for water but not for clouds. The pixels remaining after the water mask was applied were counted and converted to units of area. Only images with < 2% cloud cover over the lake were used in this estimation to avoid the classification of clouds as lake water. The number of images that were used to estimate lake surface area was 215. Images from Landsat 7 after May 2003 were excluded due to the high percentage of missing pixels from the scan line corrector error. Temporal trends in mean annual surface area and mean annual NDVI (April-October) were compared using linear models in R Studio.

RESULTS

The aims of this paper are 1) to develop a method to study changes in algae growth over time in shallow, freshwater lakes using the Landsat archive, 2) apply the method to understand the spatiotemporal patterns of algae growth on Utah Lake between 1984 and 2018, and 3) evaluate the relationship between bloom intensity, water level, and water temperature for Utah Lake. The

basic steps used in this study include compiling imagery within the seasonal date range, filtering images for cloud cover, creating a water mask, and performing hot spot analysis.

Cloud Filter Comparison

Because of the temporally dynamic nature of cyanobacteria blooms, accuracy of an historical analysis from satellite imagery decreases when images are rejected due to cloud cover. We sought an approach that maximized useable images for bloom detection while also minimizing misinterpretation due to clouds. We compared the number of useable images when the filter was done based on clouds in the entire scene or just clouds directly over Utah Lake. Filtering images by cloud cover over the lake resulted in a greater total of useable images in the analysis than filtering based on Scene Cloud Cover. Excluding images with a 10% cloud threshold over the lake retained 367 images, whereas sorting by scene cloud cover retained only 299 images (Table 2). The 5% lake cloud cover threshold resulted in 48 more total images in the analysis compared to the 10% scene cloud cover threshold. Accounting for the scan line corrector error added 110 additional images when using a 5% lake cloud cover threshold and 13 additional images when using a 20% lake cloud cover threshold (Table 3). Empty pixels due to the scan line corrector failure comprise 10.6 % of the scene.

Comparison of Four Water Masks

The SWIR1 band more accurately detected the waterbody than the NIR band under algal bloom conditions. The Modified Normalized Difference Water Index (MNDWI) uses a middle infrared band (SWIR1) instead of the near infrared band (Xu 2006). A one-way ANOVA and a Tukey's HSD Test comparing the mean number of pixels classified as water for 22 images

showed a statistically significant difference between the means of the NDWI and the MNDWI groups (one-way ANOVA, F = 2.711, $p = 0.0501$, $df = 3$; Tukey's HSD Test, NDWI-MNDWI, $p = 0.03$). The ANOVA assumptions of normality, independence, and homoscedasticity were met (Levene's Test, F = 0.3465, $p = 0.7917$, $df = 3$). The NIR and NDWI masks excluded areas of dense algae on the lake (Figures 3B, 3C, 3D). The MNDWI water mask performed the best because it distinguished land from water without excluding dense algal blooms inside the lake (Figure 3E). Visual inspection clearly revealed that the threshold for NIR values was too low and classified dense algae within the lake as if it were land (Figure 3C). Increasing the NIR threshold included significantly more land, which was problematic for future hot spot analysis. The NDWI mask showed a similar result and removed much of the lake when it was covered in algae (Figure 3B). The combined NDWI and NIR mask retained more of the lake surface area, but still removed many important areas of algae (Figure 3D). The Modified Normalized Difference Water Index (MNDWI) uses a middle infrared band (SWIR1) instead of the near infrared band (Xu 2006) and overcomes the issue of excluding dense algae without including land.

Hot Spot Parameters

One of the aims of this study was to use spatial statistics to determine areas where algal growth is increasing or decreasing on Utah Lake over time. To do this, we evaluated the results of changing the time-step in the Emerging Hot Spot Analysis. We compared the results of using 1-, 4-, 9- and 19-year neighborhood time steps, which create analysis neighborhoods of 2, 5, 10, and 20 years, respectively (Figure 4). The map with a neighborhood time step of 19 years comprises 15.3% intensifying hot spots and 0.45% diminishing hot spots (Figure 5D). The map of a neighborhood time step of 1 year is composed of 4.73% intensifying hot spots and 3.57%

diminishing hot spots (Figure 4A). Intensifying hot spots make up 5.85% and 7.62% of the maps using neighborhood time steps of 4 and 9 years, respectively (Figure 4B, 4C). A neighborhood time step of 9 results in persistent hot spots comprising 6.62% of bins, compared to 4.30% in the map with the 1-year neighborhood time step (4A, 4C).

Emerging Hot Spot Analysis

The Emerging Hot Spot Analysis shows that the majority of the lake is a cold spot, but the eastern and southern shoreline and the Provo Bay are statistically significant hot spots (Figures 5, 6, 7). The spatial pattern of hot and cold spots is similar between the results using the chlorophyll-a predictions (Figure 5), NDVI (Figure 6), and FAI (Figure 7). The Provo Bay and the eastern shoreline are intensifying and enduring hot spots. Intensifying and enduring hot spots were hot spots for 90% of the yearly time steps. In other words, the Provo Bay and the eastern shoreline were statistically significant hot spots for 30 of the 34 years included in this study. The NDVI and FAI models designated more hot spots on the western shoreline than the chl-a model. The percent match between the NDVI, FAI, and chl-a maps (Figure 8) is 81.6%, with the largest discrepancy involving the areas with no pattern detected.

Comparison of Utah Lake Water Level, NDVI, and Temperature

NDVI in hot spot areas is correlated positively and negatively to water temperature and water level, respectively. There is an inverse relationship between mean NDVI and lake surface area within hot spot areas for the 215 cloudless dates where lake surface area data were acquired (Figures 9, 10; $F = 72.63$, $R^2 = 0.254$, $p = 2.87 \times 10^{-15}$, $df = 213$). A linear model between daily mean thermal and NDVI values inside hot spot areas for all dates shows a positive relationship

(Figure 11; F =81.82, $R^2 = 0.17$, $p = 2.2 \times 10^{-16}$, $df = 396$). For 1984-2018, the overall trend of water level, and NDVI is decreasing (Figure 10, Appendix I). The Create Time Space Cubes by Aggregating Points tool uses the Mann-Kendall statistic to measure temporal trends, regardless of space. The output of this tool indicates that the overall trend of chl-a, NDVI, and FAI is decreasing ($p = 0.005$, $p = 0.0002$, and $p = 0.0000$; Appendix I, II), while lake temperature overall is increasing ($p = 0.006$; Appendix III).

DISCUSSION

The goals of this research were to develop a method to study changes in algae growth over time using the Landsat archive and to apply it to conditions on Utah Lake between 1984 and 2018.The method we used involved finding the best way to filter cloudy images, separate water from land, and find hot spots using spatial analyses.

Filtering Images by Clouds Over the Study Site Results in More Images

Although we used the best cloud masking algorithms available with Landsat data products (Foga et al. 2017), we found that filtering images by Scene Cloud Cover removed many useful images because they contained clouds in areas outside of the Utah Lake study area. There were only 85 cloud free scenes out of the 398 available images, which is too few for a spatiotemporal analysis of algal blooms. Thus, we explored other means of excluding only those images from our analysis that had clouds directly over the Utah Lake study area. This is especially important because algal blooms are sporadic and the temporal resolution of Landsat satellites is coarse. Thus, each additional image included in the analysis increases the probability of detecting a bloom. To increase the number of images included in the analysis, we identified clouds and

cloud shadows directly over the lake by counting the pixels remaining after applying the cloud and cloud shadow masks. This yielded 116 cloud-free images, which was still too few and excluded images with known algal blooms. As a result, we evaluated the number of useable images as a function of the percentage of clouds within the region of interest directly above the surface of Utah Lake. This determination required accounting for errors such as the Scan Line Corrector error that occurred in Landsat 7 in 2003. The number of useable images increased dramatically with the change in the cloud cover threshold. We chose a lake cloud cover threshold of 20%, which resulted in 398 useable images while keeping error from detecting cloud shadows at a minimum (Figure 12).

Modified Normalized Difference Water Index Is Better Under Algal Bloom Conditions

Because Utah Lake varies in water level within and between years, we needed a protocol to identify the lake surface area for each satellite image. We tested the Normalized Difference Water Index (NDWI), but determined that this approach fails to classify lake areas under algal bloom conditions as water (Table 4, Figure 3). Combinations of the NDWI and the near infrared (NIR) band also remove areas of dense algae in the middle of the lake, whereas the Modified Normalized Difference Water Index (MNDWI) does not (Figure 3D, 3E). The NIR band as a mask also failed on its own in this study (Figure 3C). The nature of the NIR is congruent with the use of the NDVI and other indices to detect photosynthetic activity in lake bodies. Thus the NIR band appears to be more sensitive to algae on the water than it is to the water itself. The effectiveness of the MNDWI has implications for use on other lakes, even under dense algal conditions. Utah Lake is unique, due to the calcium precipitates that constitute the water column. This chemistry alters the brightness and “greenness” of the lake, which differs from many other

freshwater lakes. A normalized index like the MNDWI addresses the concerns of unique chemistry and brightness and supports the use of this index on other lakes (Ho et al. 2017; Oyama et al. 2015; Xu 2006).

Neighborhood Time Step Affects the Proportion of Persistent Spots

We found that changing the neighborhood time-step in the Emerging Hot Spot Analysis modifies the resulting categories (Figure 4). A neighborhood time step of 1 year means that the dark blue square in Figure 2 is analyzed with the light blue squares for any particular year and the previous year. All values in the analysis neighborhood are summed and then compared to the sum of the entire dataset. Larger neighborhood time steps are more likely to have more persistent and intensifying hot and cold spots because each neighborhood is similar to the neighborhood in the previous year. A neighborhood time step of 1 year includes a total of 2 time steps in the neighborhood and reduces the number of persistent hot spots because it reduces the redundancy of the analysis neighborhoods. We chose a neighborhood time step of 4 because 5-year analysis neighborhoods reduce sensitivity to sporadic blooms. Harris et al. (2017) used 1-year neighborhood time steps study changes in forest loss, but we chose a larger neighborhood to average NDVI over a longer period of time.

Provo Bay and Eastern Utah Lake Are Hot Spots

The Provo Bay and the eastern shoreline are enduring or intensifying hot spots. These locations are defined by being statistically significant hot spots for 90% of the time-step intervals. The Provo Bay is distinct from the rest of the lake because it is smaller, more enclosed and does not circulate well. Most of the lake is often mixed by wind, whereas the Provo Bay is

more isolated from these wind patterns. The hot spots on the eastern side of the lake correspond with a large and growing urban population and are characterized by high phosphorus concentrations (Randall et al. 2019). The dominance of the cold spot on Utah Lake suggests that algal blooms on the lake as a whole are not worse than they were in the mid 1980's (Figures 5, 6, 7). The no pattern space between the hot and cold spots in the hot spot results of the chlorophyll-a, NDVI, and Floating Algae Index (FAI) models (Figures 5, 6, 7) represents transition zones and areas of algal diffusion through water. These three hot spot maps show similar patterns, with an 81.6% rate of matching (Figure 8). The similarities between the results of the three models have implications for scientists studying other lakes as well as using imagery from earlier Landsat satellites, which are characterized by a low spectral resolution. The NDVI can be applied to other lakes and satellite imagery lacking short-wave infrared bands.

Relationship Between NDVI in Hot Spots, Lake Water Level, and Water Temperature

Algal concentration in hot spots intensifies with declining water level and increasing temperature (Figures 9, 11). Declining lake surface area has implications for a population that is projected to grow at a rate of 1.3% annually, exceeding national growth rates (Perlich et al. 2017). The decrease in water level has an indirect effect on algal growth. Lower water levels may lead to increases in temperature and nutrient concentrations, providing ideal conditions for algal growth. Utah Lake's water temperature shows an upward trend and is an important consideration for a changing climate ($p = 0.006$; Appendix III). The positive correlation between NDVI and water temperature supports the hypothesis that NDVI is higher in warmer temperatures, as was found in other studies (Walls et al. 2018). Although NDVI in the hot spots is increasing, mean NDVI over time is decreasing, supporting the hypothesis that Utah Lake's

water quality has not declined between 1984 and 2018 (Figure 10). Contrastingly, Hansen et al. (2019) found an increasing trend for both mean and maximum annual chl-a. Key differences between our study and Hansen et al. (2019) include the use of the entire lake to predict NDVI and chl-a, rather than the average of discrete sampling points. Due to the size and spatial variation in Utah Lake, not accounting for the entire lake could exclude important information. The relationships between NDVI, temperature, and water level provide valuable insights into the past and have implications for a changing climate and a growing population. The methods of our study can be used to understand the spatiotemporal dynamics of algal growth on lakes worldwide. Scientists can use these methods to compile historical records of lakes larger than the Landsat spatial resolution to better predict and manage algal blooms in the future.

REFERENCES

Augusto-Silva, P.B., Ogashawara, I., Barbosa, C.C.F., de Carvalho, L.A.S., Jorge, D.S.F., Fornari, C.I., & Stech, J.L. (2014). Analysis of MERIS Reflectance Algorithms for Estimating Chlorophyll-a Concentration in a Brazilian Reservoir. *Remote Sensing, 6*, 11689-11707

Bertani, I., Steger, C.E., Obenour, D.R., Fahnenstiel, G.L., Bridgeman, T.B., Johengen, T.H., Sayers, M.J., Shuchman, R.A., & Scavia, D. (2017). Tracking cyanobacteria blooms: Do different monitoring approaches tell the same story? *Science of the Total Environment, 575*, 294-308

Brezonik, P., Menken, K.D., & Bauer, M. (2005). Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (CDOM). *Lake and Reservoir Management, 21*, 373-382

Clark, J.M., Schaeffer, B.A., Darling, J.A., Urquhart, E.A., Johnston, J.M., Ignatius, A.R., Myer, M.H., Loftin, K.A., Werdell, P.J., & Stumpf, R.P. (2017). Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources. *Ecological Indicators, 80*, 84-95

da Costa, M.R.A., Attayde, J.L., & Becker, V. (2016). Effects of water level reduction on the dynamics of phytoplankton functional groups in tropical semi-arid shallow lakes. *Hydrobiologia, 778*, 75-89

DEQ (2014). Utah Lake Algal Bloom October 2014. In. deq.utah.gov: Utah Department of Environmental Quality

DEQ (2016). Utah Lake, Jordan River, Canals Algal Bloom 2016. In. deq.utah.gov: Utah Department of Environmental Quality

DEQ (2017). Utah Lake, Jordan River, Canals Algal Bloom 2017. In: Utah Department of Environmental Quality

DEQ (2018). Utah Lake Algal Bloom Monitoring 2018. In U.D.o.W. Quality (Ed.), *2018 Monitoring Updates*. deq.utah.gov: Utah Department of Environmental Quality

DWR, U. (2014). Utah Lake Basin Planning for the Future In, *Utah State Water Plan* (pp. 1-143). water.utah.gov: Utah Division of Water Resources

EPA, U.S. (2018). Cyanobacteria/Cyanotoxins. In. epa.gov: EPA

Esri (2016). How Emerging Hot Spot Analysis Works. In, *ArcMap*. desktop.arcgis.com: Environmental Systems Research Institue, Inc.

Esri (2018a). Create Space Time C ube By Aggregating Points. In, *Space Time Pattern Mining toolbox*. pro.arcgis.com: Esri

Esri (2018b). How Creating a Space Time Cube works. In, *Space Time Pattern Mining Concepts*. pro.arcgis.com: Esri

Esri (2018c). How Hot Spot Analysis (Getis-Ord Gi*) works. In, *ArcMap*. desktop.arcgis.com: Esri

Esri (2018d). What is a z-score? What is a p-value? In, *Spatial Statistics toolbox*. pro.arcgis.com: Esri

Fadel, A., Atoui, A., Lemaire, B.J., Vincon-Leite, B., & Slim, K. (2015). Environmental factors associated with phytoplankton succession in a Mediterranean reservoir with a highly fluctuating water level. *Environmental Monitoring and Assessment, 187*

Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Hughes, M.J., & Laue, B. (2017). Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sensing of Environment, 194*, 379-390

Google (2017). Landsat Algorithms. In. developers.google.com: Google

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*, 18-27

Hansen, C., Burian, S., Dennison, P., & Williams, G. (2019). Constructing a Historical Record of Algal Blooms and their Response to Local Climate Conditions using Google Earth Engine and Remote Sensing. In: University of Utah

Harris, N.L., Goldman, E., Gabris, C., Nordling, J., Minnemeyer, S., Ansari, S., Lippmann, M., Bennett, L., Raad, M., Hansen, M., & Potapov, P. (2017). Using spatial statistics to identify emerging hot spots of forest loss. *Environmental Research Letters, 12*, 13

Ho, J.C., Stumpf, R.P., Bridgeman, T.B., & Michalak, A.M. (2017). Using Landsat to extend the historical record of lacustrine phytoplankton blooms: A Lake Erie case study. *Remote Sensing of Environment, 191*, 273-285

Hu, C.M. (2009). A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment, 113*, 2118-2129

Kloiber, S.N., Brezonik, P.L., Olmanson, L.G., & Bauer, M.E. (2002). A procedure for regional lake water clarity assessment using Landsat multispectral data. *Remote Sensing of Environment, 82*, 38-47

Mann, H.B. (1945). NONPARAMETRIC TESTS AGAINST TREND. *Econometrica, 13*, 245-259

McFeeters, S.K. (1996). The use of the normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing, 17*, 1425-1432

Miller, S.A., & Crowl, T.A. (2006). Effects of common carp (Cyprinus carpio) on macrophytes and invertebrate communities in a shallow lake. *Freshwater Biology, 51*, 85-94

Noack, U., Herden, N., Loffler, J., Warcup, C., & Gorsler, M. (1985). CONTINUOUS MEASUREMENT OF ALGAE BIOMASS BY MEANS OF CHLOROPHYLL FLUORESCENCE IN MONITORING STATIONS IN LOWER SAXONY. *Zeitschrift Fur Wasser Und Abwasser Forschung-Journal for Water and Wastewater Research, 18*, 177-182

Ogashawara, I., Li, L., & Moreno-Madrinan, M.J. (2016). Slope algorithm to map algal blooms in inland waters for Landsat 8/Operational Land Imager images. *Journal of Applied Remote Sensing, 11*, 18

Ord, J.K., & Getis, A. (1995). LOCAL SPATIAL AUTOCORRELATION STATISTICS - DISTRIBUTIONAL ISSUES AND AN APPLICATION. *Geographical Analysis, 27*, 286-306

Oyama, Y., Matsushita, B., & Fukushima, T. (2015). Distinguishing surface cyanobacterial blooms and aquatic macrophytes using Landsat/TM and ETM plus shortwave infrared bands. *Remote Sensing of Environment, 157*, 35-47

Paerl, H.W., & Paul, V.J. (2012). Climate change: Links to global expansion of harmful cyanobacteria. *Water Research, 46*, 1349-1363

Page, B.P., Kumar, A., & Mishra, D.R. (2018). A novel cross-satellite based assessment of the spatio-temporal development of a cyanobacterial hatinful algal bloom. *International Journal of Applied Earth Observation and Geoinformation, 66*, 69-81

Perlich, P.S., Hollingshaus, M., Harris, E.R., Tennert, J., & Hogue, M.T. (2017). Utah's Long-Term Demographic and Economic Projections Summary. In (pp. 1-32). Salt Lake City, Utah: The University of Utah

PSOMAS (2007). Utah Lake TMDL: Pollutant Loading Assessment & Designated Beneficial Use Impairment Assessment. In (p. 88): State of Utah Division of Water Quality

Randall, M.C., Carling, G.T., Dastrup, D.B., Miller, T., Nelson, S.T., Rey, K.A., Hansen, N.C., Bickmore, B.R., & Aanderud, Z.T. (2019). Sediment potentially controls in-lake phosphorus cycling and harmful cyanobacteria in shallow, eutrophic Utah Lake. *Plos One, 14*, 17

Sadeghian, A., Chapra, S.C., Hudson, J., Wheater, H., & Lindenschmidt, K.E. (2018). Improving in-lake water quality modeling using variable chlorophyll a/algal biomass ratios. *Environmental Modelling & Software, 101*, 73-85

Schmidt, G., Jenkerson, C., Masek, J., Vermote, E., & Gao, F. (2013). Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) Algorithm Description. In. Reston, Virginia: U.S. Geological Survey

Strong, A. (1974). Remote sensing of algal blooms by aircraft and satellite in Lake Erie and Utah Lake. *Remote Sensing of Environment, 3*, 99-107

Stumpf, R.P., Davis, T.W., Wynne, T.T., Graham, J.L., Loftin, K.A., Johengen, T.H., Gossiaux, D., Palladino, D., & Burtner, A. (2016). Challenges for mapping cyanotoxin patterns from remote sensing of cyanobacteria. *Harmful Algae, 54*, 160-173

Taranu, Z.E., Gregory-Eaves, I., Leavitt, P.R., Bunting, L., Buchaca, T., Catalan, J., Domaizon, I., Guilizzoni, P., Lami, A., McGowan, S., Moorhouse, H., Morabito, G., Pick, F.R., Stevenson, M.A., Thompson, P.L., & Vinebrooke, R.D. (2015). Acceleration of cyanobacterial dominance in north temperate-subarctic lakes during the Anthropocene. *Ecology Letters, 18*, 375-384

Tebbs, E.J., Remedios, J.J., & Harper, D.M. (2013). Remote sensing of chlorophyll-a as a measure of cyanobacterial biomass in Lake Bogoria, a hypertrophic, saline-alkaline, flamingo lake, using Landsat ETM. *Remote Sensing of Environment, 135*, 92-106

Urquhart, E.A., Schaeffer, B.A., Stumpf, R.P., Loftin, K.A., & Werdell, P.J. (2017). A method for examining temporal changes in cyanobacterial harmful algal bloom spatial extent using satellite remote sensing. *Harmful Algae, 67*, 144-152

USGS (2017). Landsat Collections. In, *Landsat Missions*. landsat.usgs.gov: United States Geological Survey

USGS (2018a). Landsat Collection 1 Level-1 Quality Assessment Band. In, *Landsat*. landsat.usgs.gov: United States Geological Survey

USGS (2018b). Question: How is the percentage of cloud cover calculated in a Landsat scene? In, *FAQs*. landsat.usgs.gov: United States Geological Survey

Walls, J.T., Wyatt, K.H., Doll, J.C., Rubenstein, E.M., & Rober, A.R. (2018). Hot and toxic: Temperature regulates microcystin release from cyanobacteria. *Science of the Total Environment, 610*, 786-795

Whiting, M.C., Brotherson, J.D., & Rushforth, S.R. (1978). Environmental interaction in summer algal communities of Utah Lake. *Great Basin Naturalist, 38*, 31-41

Xu, H.Q. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing, 27*, 3025-3033

FIGURES

Figure 1. The small box on the inset map of the United States shows the location of Utah Lake. The expanded map of Utah Lake shows the rivers that enter the lake in blue and wastewater treatment plant inputs as green triangles, all located in the urban areas on the East side of the lake. Three wastewater treatment plants empty into the Provo Bay, which is the area where Hobble Creek empties into Utah Lake.

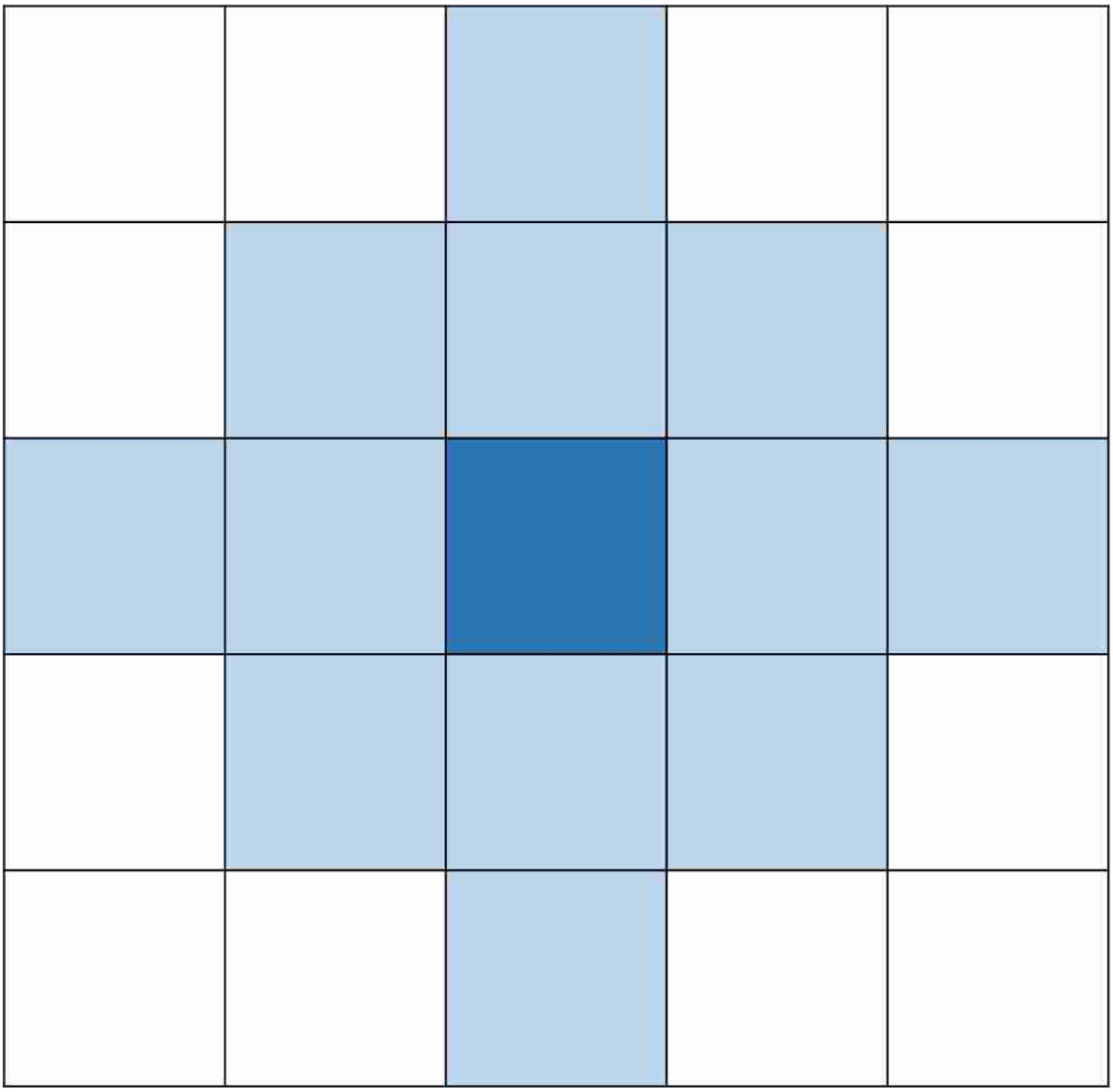

Figure 2. This figure illustrates the spatial sampling approach used in the hot spot analysis, which uses the Queens Case Contiguity method. The dark blue square in the center represents each bin in the hot spot analysis, whereas the surrounding light blue squares represent the bins that it is compared to in order to find clusters of high or low values. High values in this neighborhood relative to the values in the rest of the time series result in a hot spot.

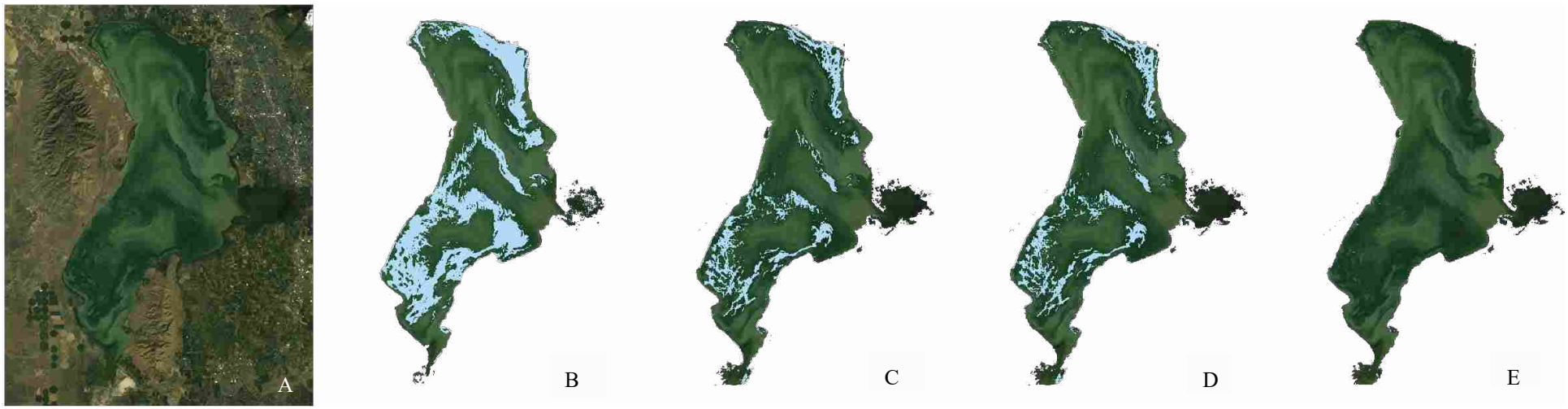

Figure 3. A comparison of four different water mask approaches evaluated using a Landsat 5 image during a known algal bloom on September 10, 2006 in the Google Earth Engine. Image A is the original image. Image B is the mask that uses the Normalized Difference Water Index (NDWI) and shows areas within the lake where high cyanobacteria concentrations cause the water mask to fail (blue shaded areas). Image C is the mask made with made with only NIR and image D used both NIR and NDWI conditions. Image E used the Modified Normalized Difference Water Index (MNDWI).

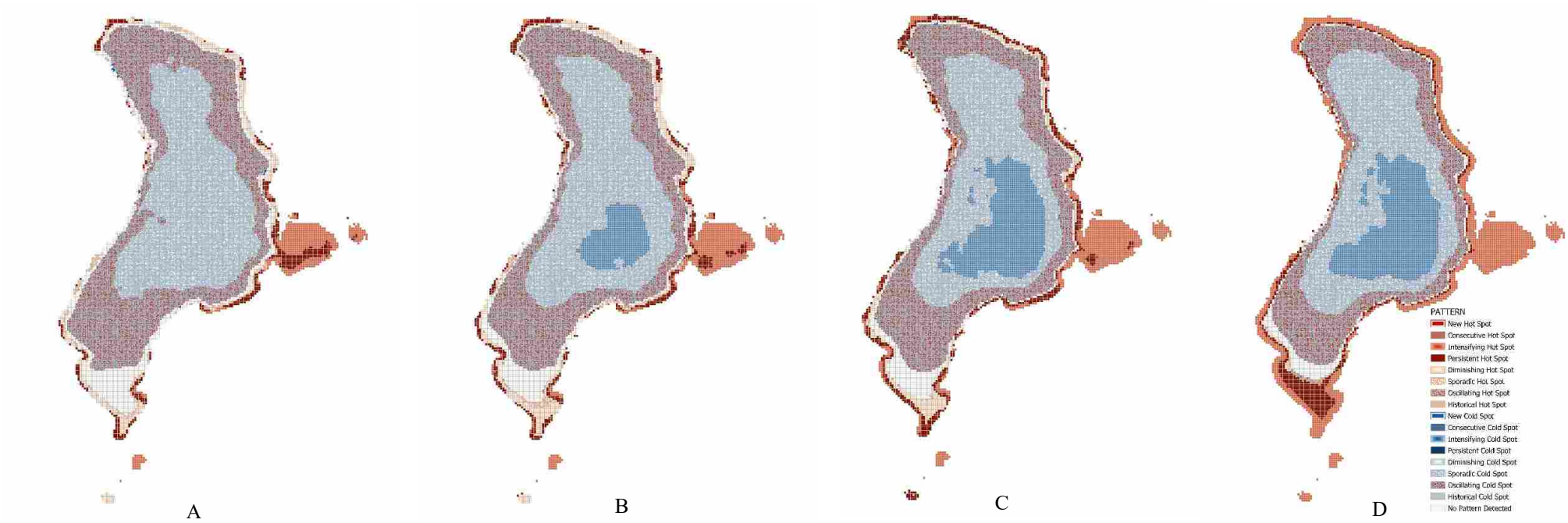

Figure 4. These maps compare the impact of various neighborhood time steps on the emerging hot spot analysis. The maps from left to right use 1- (A), 4- (B), 9- (C), and 19- (D) year neighborhood time steps. The larger neighborhood time steps are related to a higher proportion of persistent and intensifying hot or cold spots. This is because larger neighborhood time steps have similar neighborhoods for each year.

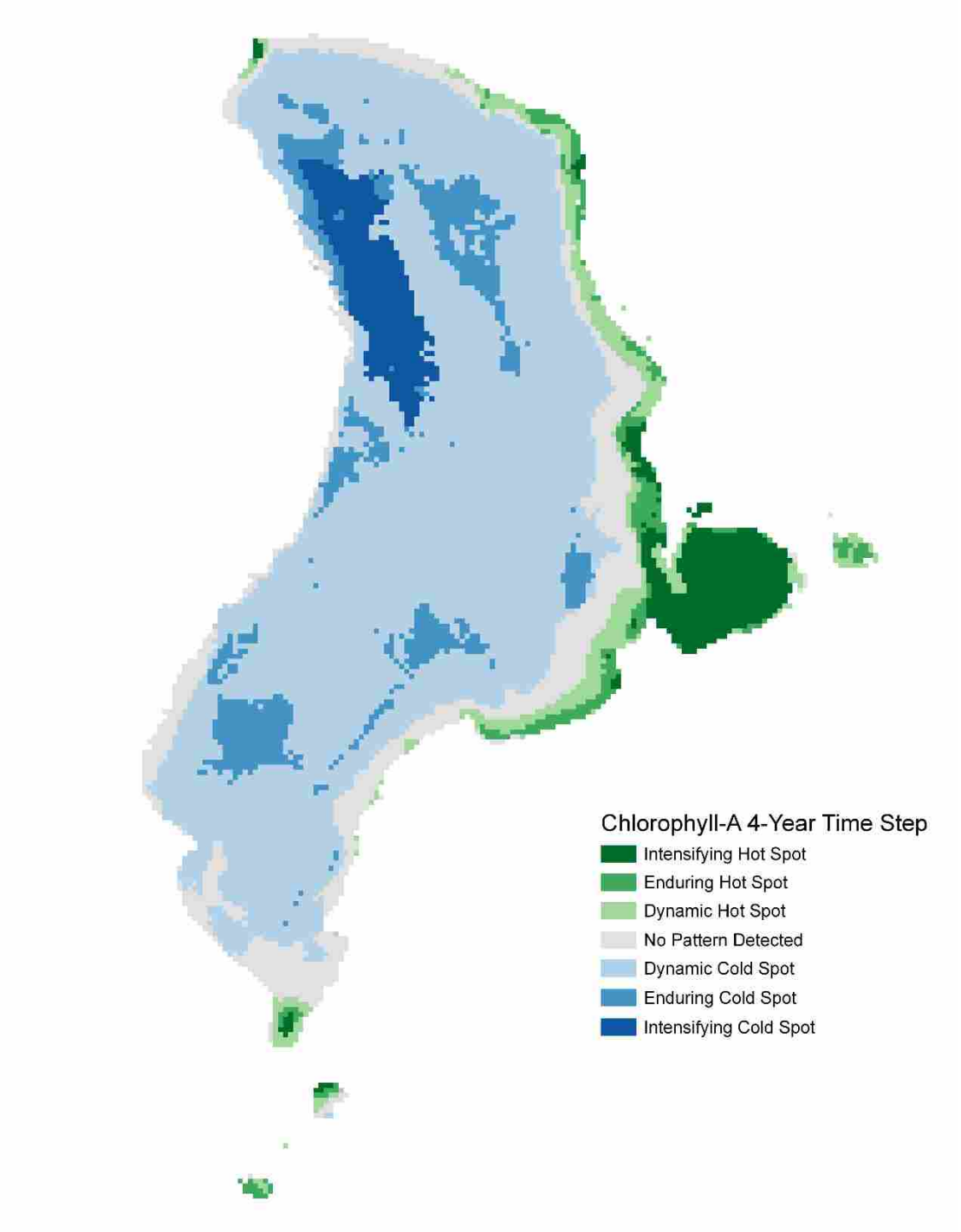

Figure 5. Results of the Emerging Hot Spot Analysis using predictions of chlorophyll-a. The majority of the lake is a cold spot, while the eastern side, including the shoreline and the Provo Bay are hot spots of algal biomass.

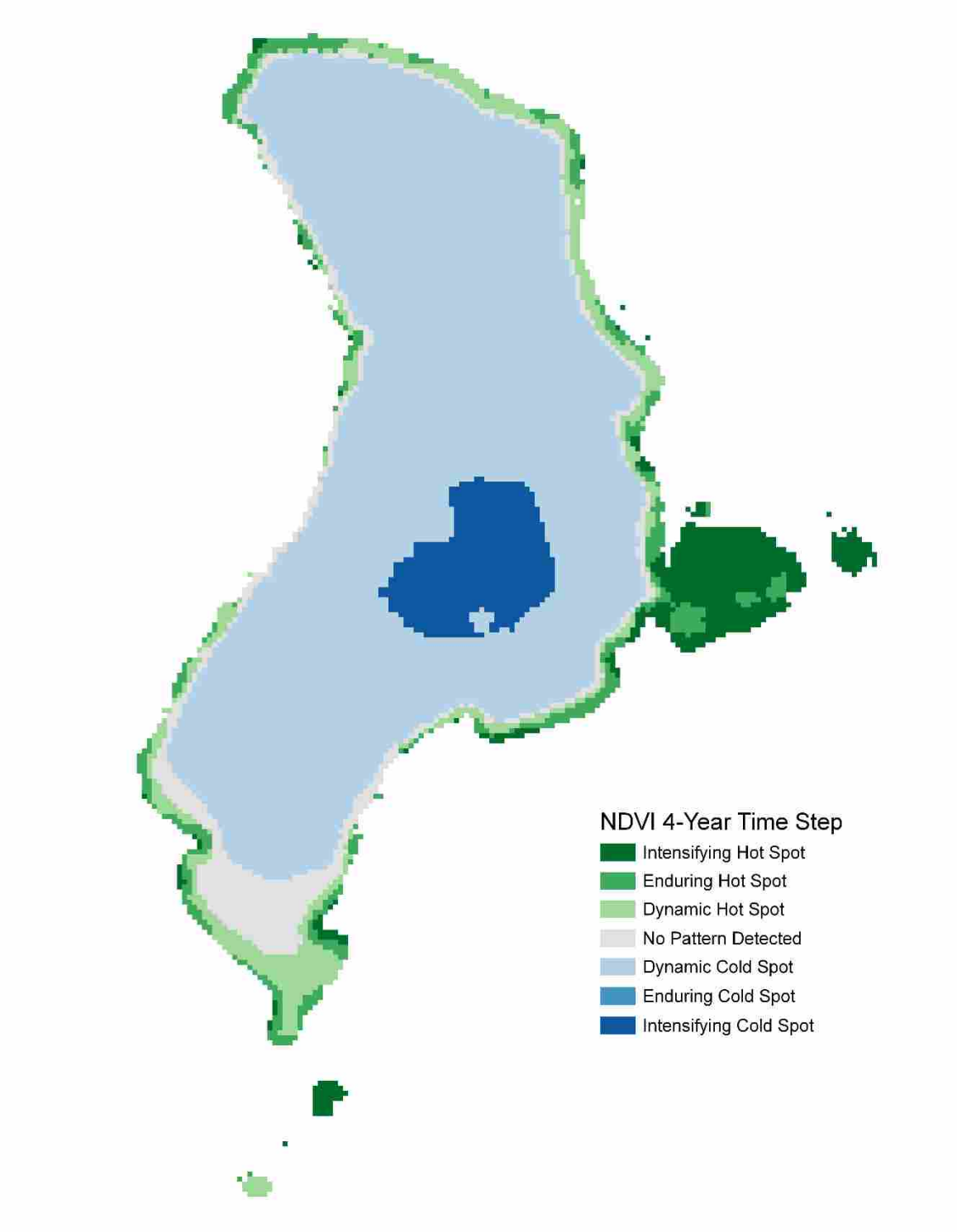

Figure 6. Results of the Emerging Hot Spot Analysis using the Normalized Difference Vegetation Index (NDVI) as a measure of photosynthetic activity. The majority of the lake is a cold spot, while the eastern side and some of the western shoreline are hot spots of algal biomass. The Provo Bay stands out as mostly an intensifying hot spot.

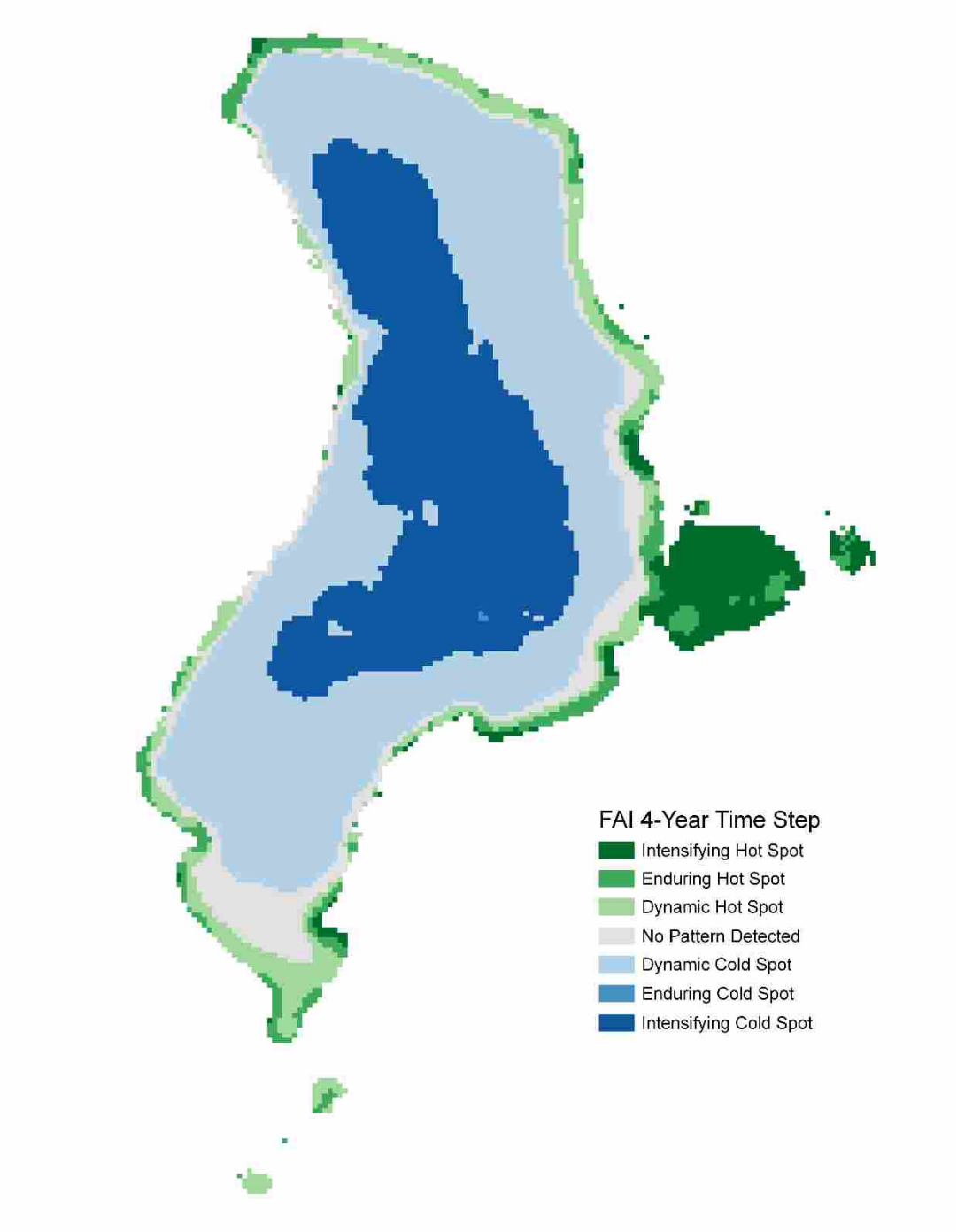

Figure 7. Results of the Emerging Hot Spot Analysis using the Floating Algae Index (FAI) as a measure of algae biomass. Most of the lake is a cold spot, while most of the shoreline and the Provo Bay are hot spots of algal biomass.

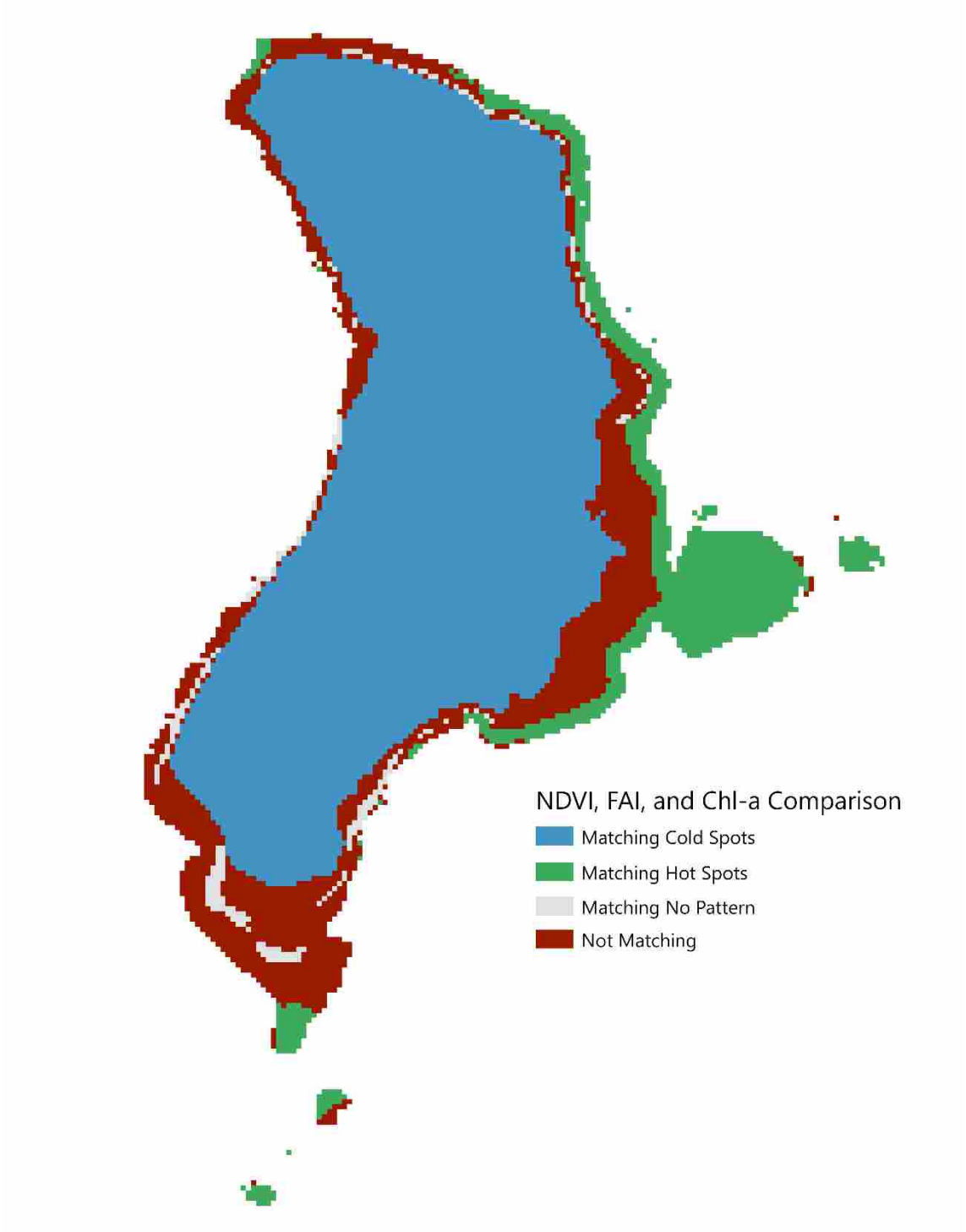

Figure 8. This map is a comparison of the results of the hot spot analysis between the algal density predictions based on the chlorophyll-a model, the Floating Algae Index (FAI), and the Normalized Difference Vegetation Index (NDVI). Bins that differed between the three outputs are shown in red. This shows that the results of the hot spot analysis using distinct models to predict algal density are similar. This has implications for the use of NDVI and the FAI to predict algal growth other lakes as well as with earlier satellites, such as Landsat 1, 2, and 3.

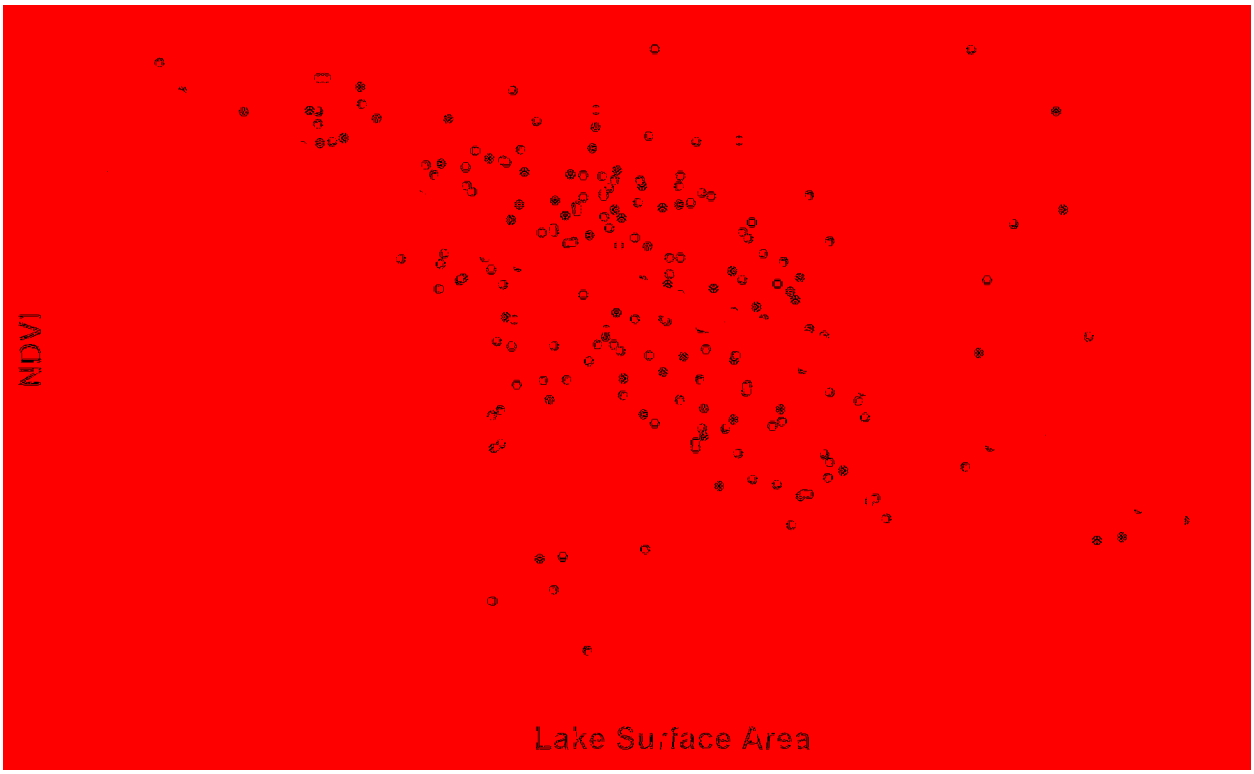

Figure 9. Relationship between lake surface area (km^2) and the mean Normalized Difference Vegetation Index (NDVI) for points within hot spot areas (Figure 8) on each date that lake surface area data were available. The number of dates included in the graph is 215, which is the total number of images with a 2% cloud cover threshold, excluding images from Landsat 7 taken after May 31, 2003. The relationship between lake surface area and NDVI is inverse (R^2 = 0.254).

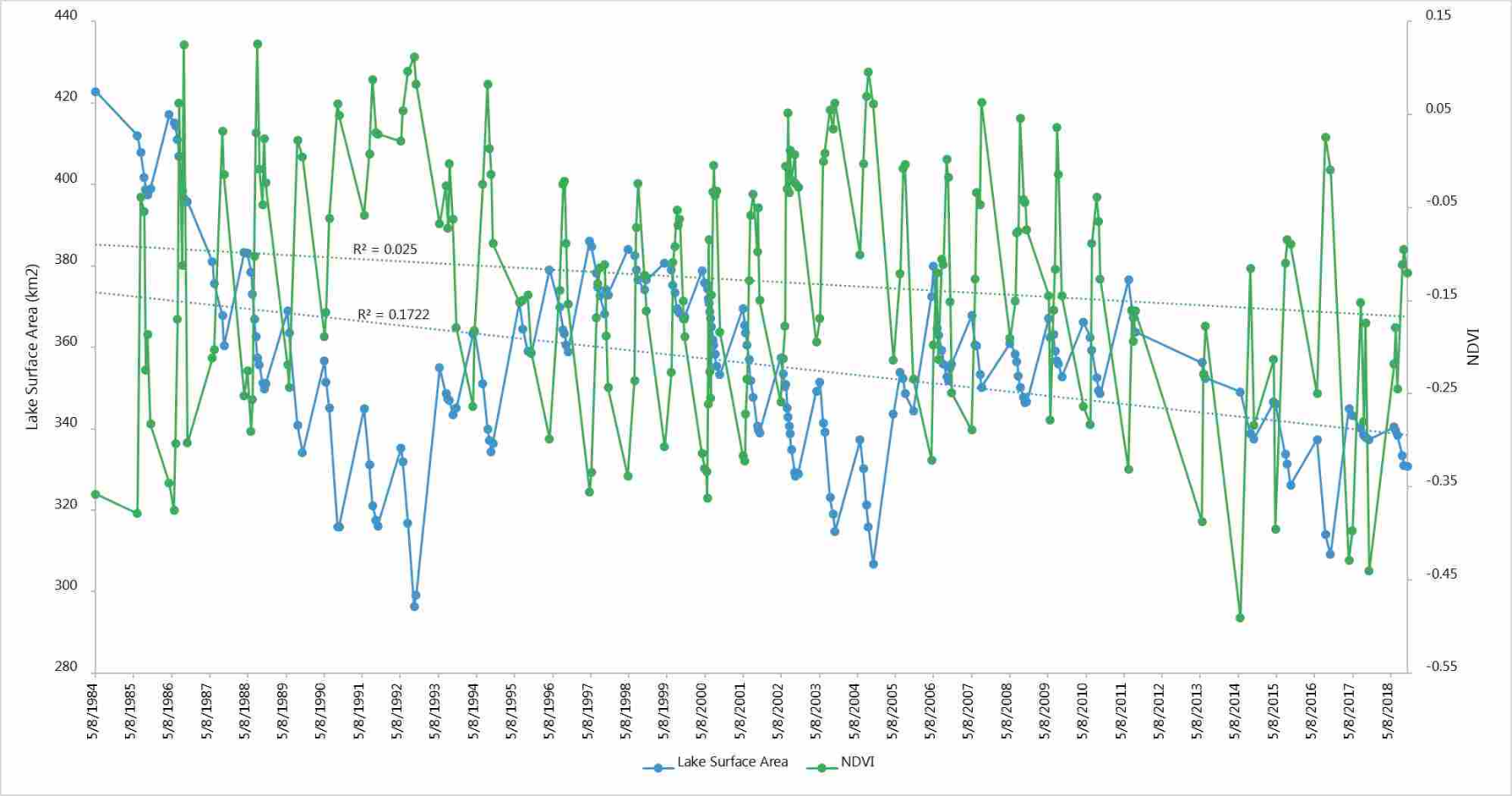

Figure 10. Time series chart of daily means in hot spot areas of lake surface area (blue) and NDVI (green). Lake surface area was used as an approximation of lake elevation estimated from images with <2% lake cloud cover. Water level varies greatly within and between years and is inversely related to mean NDVI.

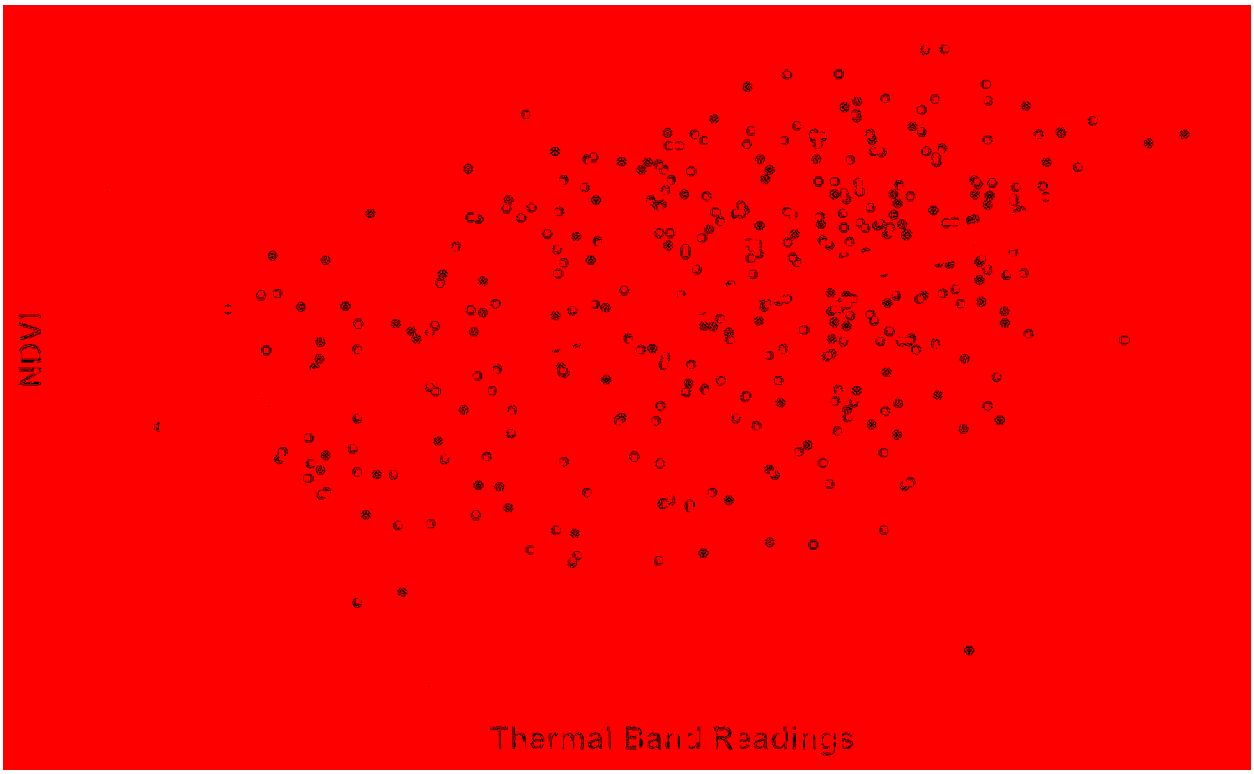

Figure 11. Relationship between the thermal band readings from the satellites and the mean Normalized Difference Vegetation Index (NDVI) for points within hot spot areas (Figure 8) on each date (398 dates). The relationship between lake temperature and NDVI is positive (R^2 = 0.171).

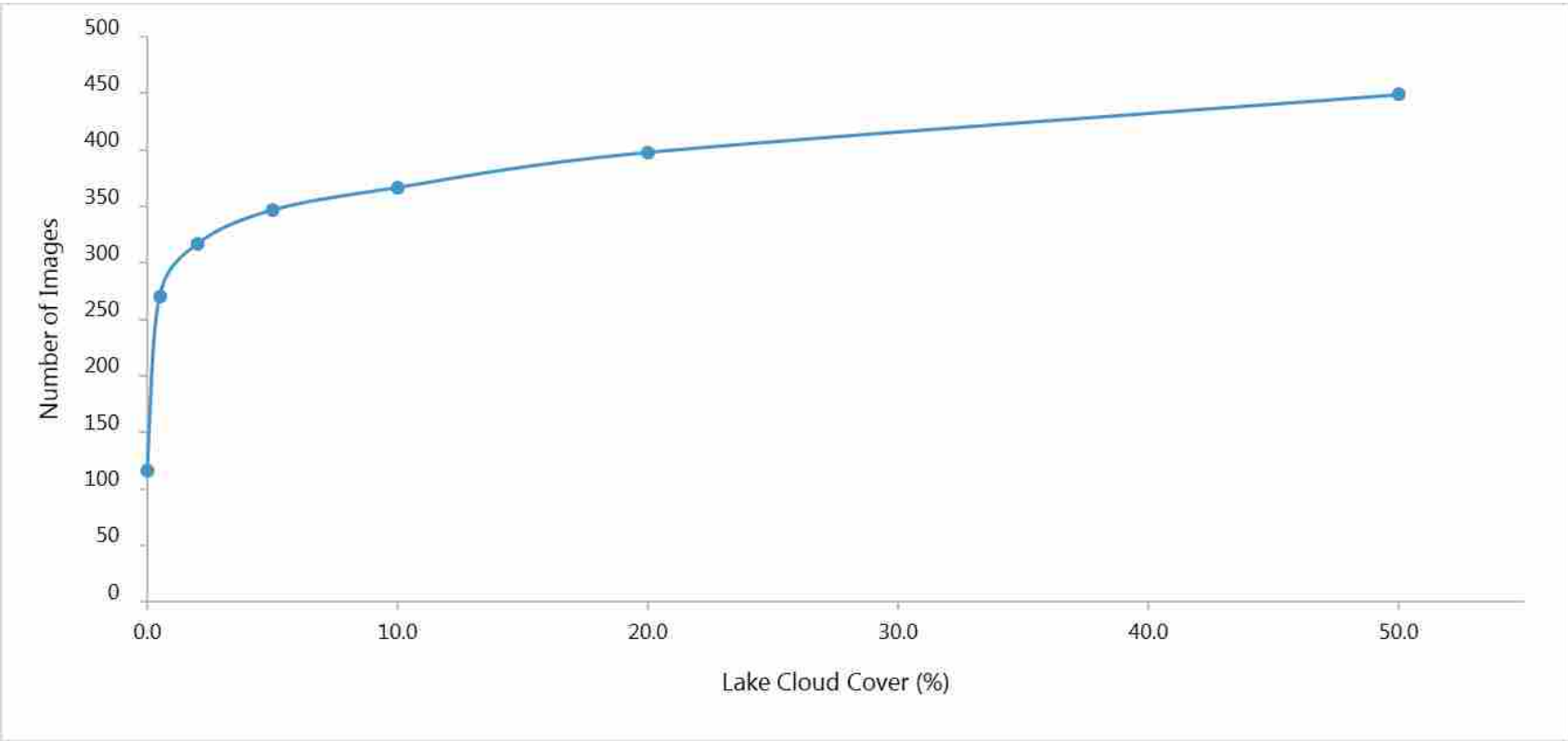

Figure 12. Total number of images that could be utilized in the hot spot analysis as a function of the variation in the threshold for percent of the lake surface obscured by clouds in the satellite images over Utah Lake. The largest increase in total useable images is between the threshold values of 0 and 5% clouds covering the lake. The figure illustrates the difficulty of obtaining cloud-free imagery. The difference between thresholds of 10 and 20% is 31 images. Because the study seeks to identify algal blooms on the lake and because algal blooms are temporally sporadic, a threshold of 20% was chosen to include more images in the analysis and more accurately detect their temporal patterns.

TABLES

Table 1. This table describes the modifications we made to the original output categories of the Emerging Hot Spot Analysis Tool (Esri 2016). The enduring and dynamic categories apply to both hot and cold spots. Enduring cold spots were cold spots for greater than 90% of the time-steps and dynamic cold spots were cold spots for less than 90% of the time-steps (years).

Original Category	**New Category**	**New Category Description**
Intensifying Hot Spot	Intensifying Hot Spot	Location where greater than 90% of the time-step intervals are statistically significant hot spots, including the last time step. There is a statistically significant increase in the clustering of high values.
Persistent	Enduring	Location where greater than 90% of the time-step intervals are statistically significant hot or cold spots.
Historical		
Diminishing		
Sporadic	Dynamic	Location where less than 90% of the time-step intervals are statistically significant hot or cold spots.
Oscillating		
New		
Consecutive		
Intensifying Cold Spot	Intensifying Cold Spot	Location where greater than 90% of the time-step intervals are statistically significant cold spots, including the last time step. There is a statistically significant increase in the clustering of low values.
No Pattern Detected	No Pattern Detected	Location that does not fit into the above categories.

Table 2. A comparison of the number of images that meet various conditions of image cloud cover. This table accounts for missing data due to Landsat 7's Scan Line Corrector (SLC) error by calculating the percentage of lake cloud cover from the total number of pixels unaffected by the SLC failure. Landsat 7 scenes before and after the SLC failure on May 31, 2003 are considered "pre" and "post," respectively. Filtering for cloud cover directly above the lake rather than within the entire scene increases the number of images in the collection, even with a stricter 5% threshold.

	No Filter	Scene Cloud Cover			Lake Cloud Cover						
		0%	<10%	<20%	0%	<0.5%	<2%	<5%	<10%	<20%	<50%
Landsat 4, 5	309	48	153	198	17	145	172	186	194	205	226
Landsat 7 Pre	48	7	21	116	0	11	17	20	21	23	30
Landsat 7 Post	191	30	95	26	98	98	102	108	113	125	144
Landsat 8	76	0	30	41	1	16	26	33	39	45	49
Total	**624**	**85**	**299**	**381**	**116**	**270**	**317**	**347**	**367**	**398**	**449**

Table 3. This table shows a comparison of total images in the collection from before and after applying the corrections due to the scan line corrector failure of Landsat 7. The correction added 108 images to the collection using the 5% threshold of clouds directly over the lake. The correction also added 74 images to the total image count with the 10% lake cloud cover threshold.

	Lake Cloud Cover						
	0%	<0.5%	<2%	<5%	<10%	<20%	<50%
Without Correction	18	172	215	239	293	385	448
With Correction	116	270	317	347	367	398	449

Table 4. The area of Utah Lake classified as water with the four water mask approaches using satellite images taken on 22 different dates. The Normalized Difference Water Index (NDWI) has the lowest mean as a result of misclassifying areas of dense algal growth within the lake as land. The Modified Normalized Difference Water Index (MNDWI) has the highest mean and does not exclude dense areas of algal growth.

Area Classified as Water (km^2)				
	NDWI with NIR	NDWI	NIR	MNDWI
5/8/1984	416	411	414	420
9/9/1985	344	305	344	395
7/1/1986	405	388	405	409
7/17/1986	387	324	387	405
9/6/1987	295	293	238	364
9/22/1987	329	304	328	358
7/6/1988	348	330	292	364
7/22/1988	351	319	351	360
8/26/1989	324	303	324	339
6/24/1995	368	349	368	367
9/12/1995	358	342	358	357
9/1/1997	345	329	345	345
9/17/1997	367	351	367	366
6/21/2000	364	356	364	366
9/7/2005	330	319	330	335
9/15/2005	303	293	303	306
9/2/2006	304	270	304	309
4/1/2017	347	336	346	343
5/3/2017	343	336	343	341
6/23/2018	337	333	337	338
8/10/2018	320	320	316	324
8/26/2018	318	313	309	328
Mean	**339**	**328**	**337**	**342**
Standard Deviation	**30.5**	**30.6**	**38.1**	**29.9**

APPENDIX I

Messages

Start Time: Thursday, April 11, 2019 4:44:59 PM

```
Running script CreateSpaceTimeCube...
The space time cube has aggregated 16452933 points into 34717 fishnet grid locations over 35 time step
intervals.  Each location is a 180 meters by 180 meters square.  The entire space time cube spans an
area 26820 meters west to east and 41940 meters north to south.  Each of the time step intervals is 1
year in duration so the entire time period covered by the space time cube is 35 years.  Of the 34717
total locations, 14180 (40.84%) contain at least one point for at least one time step interval.  These
14180 locations comprise 496300 space time bins of which 417061 (84.03%) have point counts greater than
zero.  There is a statistically significant increase in point counts over time.
---------- Space Time Cube Characteristics -----------
Input feature time extent          1984-05-08 00:00:00
                                to 2018-10-13 00:00:00

Number of time steps                                35
Time step interval                           1 year
Time step alignment                             End

First time step temporal bias                  51.64%
First time step interval                        after
                                    1983-11-01 00:00:00
                                        to on or before
                                    1984-11-01 00:00:00

Last time step temporal bias                    5.21%
Last time step interval                         after
                                    2017-11-01 00:00:00
                                        to on or before
                                    2018-11-01 00:00:00

Cube extent across space       (coordinates in meters)
Min X                                     419415.0000
Min Y                                    4426965.0000
Max X                                     446235.0000
Max Y                                    4468905.0000
Rows                                             233
Columns                                          149
Total bins                                   1215095
----------------------- COUNT ------------------------
Total number of locations                      34717
Locations with at least one point              14180
- associated bins                             496300
- % non-zero (sparseness)                      84.03

---- Summary Field - NDVI_MEAN_SPATIAL_NEIGHBORS -----
% of locations excluded due to unfilled bins     7.52%
- Total number                                    2611
Total number of locations                      11569
- associated bins                              404915
% of included locations with estimated bins      8.71%
- Total number                                    1008
% of all bins that were estimated                3.68%
- Total number                                   14916

----- Summary Field - FAI_MEAN_SPATIAL_NEIGHBORS -----
% of locations excluded due to unfilled bins     7.52%
- Total number                                    2611
Total number of locations                      11569
- associated bins                              404915
% of included locations with estimated bins      8.71%
- Total number                                    1008
% of all bins that were estimated                3.68%
- Total number                                   14916

----- Summary Field - NIR_MEAN_SPATIAL_NEIGHBORS -----
```

```
  % of locations excluded due to unfilled bins     7.52%
  - Total number                                    2611
  Total number of locations                        11569
  - associated bins                               404915
  % of included locations with estimated bins      8.71%
  - Total number                                    1008
  % of all bins that were estimated                3.68%
  - Total number                                   14916

  ------------- Overall Data Trend - COUNT -------------
  Trend direction                              Increasing
  Trend statistic                                  4.3172
  Trend p-value                                    0.0000
  -- Overall Data Trend - NDVI_MEAN_SPATIAL_NEIGHBORS --
  Trend direction                              Decreasing
  Trend statistic                                 -3.7492
  Trend p-value                                    0.0002
  -- Overall Data Trend - FAI_MEAN_SPATIAL_NEIGHBORS ---
  Trend direction                              Decreasing
  Trend statistic                                 -4.1468
  Trend p-value                                    0.0000
  -- Overall Data Trend - NIR_MEAN_SPATIAL_NEIGHBORS ---
  Trend direction                              Decreasing
  Trend statistic                                 -2.6699
  Trend p-value                                    0.0076
  Completed script Create Space Time Cube By Aggregating Points...
Succeeded at Thursday, April 11, 2019 4:50:30 PM (Elapsed Time: 5 minutes 30 seconds)
```

APPENDIX II

Messages

Start Time: Thursday, April 11, 2019 5:14:48 PM

```
Running script CreateSpaceTimeCube...
The space time cube has aggregated 15143978 points into 34484 fishnet grid locations over 35 time step
intervals.  Each location is a 180 meters by 180 meters square.  The entire space time cube spans an
area 26640 meters west to east and 41940 meters north to south.  Each of the time step intervals is 1
year in duration so the entire time period covered by the space time cube is 35 years.  Of the 34484
total locations, 14092 (40.87%) contain at least one point for at least one time step interval.  These
14092 locations comprise 493220 space time bins of which 416173 (84.38%) have point counts greater than
zero.  There is a statistically significant increase in point counts over time.

---------- Space Time Cube Characteristics -----------
Input feature time extent          1984-05-08 00:00:00
                                to 2018-10-13 00:00:00

Number of time steps                                 35
Time step interval                             1 year
Time step alignment                               End

First time step temporal bias                    51.64%
First time step interval                          after
                                   1983-11-01 00:00:00
                                       to on or before
                                   1984-11-01 00:00:00

Last time step temporal bias                      5.21%
Last time step interval                           after
                                   2017-11-01 00:00:00
                                       to on or before
                                   2018-11-01 00:00:00

Cube extent across space       (coordinates in meters)
Min X                                      419595.0000
Min Y                                     4426965.0000
Max X                                      446235.0000
Max Y                                     4468905.0000
Rows                                              233
Columns                                           148
Total bins                                    1206940

----------------------- COUNT ------------------------
Total number of locations                        34484
Locations with at least one point                14092
- associated bins                               493220
- % non-zero (sparseness)                         84.38
 Summary Field - CHLOROPHYLL_A_MEAN_SPATIAL_NEIGHBORS
% of locations excluded due to unfilled bins      7.43%
- Total number                                    2563
Total number of locations                        11529
- associated bins                               403515
% of included locations with estimated bins       8.49%
- Total number                                     979
% of all bins that were estimated                3.51%
- Total number                                   14178

------------- Overall Data Trend - COUNT -------------
Trend direction                              Increasing
Trend statistic                                  4.5445
Trend p-value                                    0.0000
 Overall Data Trend -
CHLOROPHYLL_A_MEAN_SPATIAL_NEIGHBORS
Trend direction                              Decreasing
Trend statistic                                 -2.8119
Trend p-value                                    0.0049
Completed script Create Space Time Cube By Aggregating Points...
```

APPENDIX III

Messages

Start Time: Thursday, April 11, 2019 5:22:18 PM

```
Running script CreateSpaceTimeCube...
The space time cube has aggregated 16452933 points into 34717 fishnet grid locations over 35
time step intervals.  Each location is a 180 meters by 180 meters square.  The entire space
time cube spans an area 26820 meters west to east and 41940 meters north to south.  Each of
the time step intervals is 1 year in duration so the entire time period covered by the space
time cube is 35 years.  Of the 34717 total locations, 14180 (40.84%) contain at least one
point for at least one time step interval.  These 14180 locations comprise 496300 space time
bins of which 417061 (84.03%) have point counts greater than zero.  There is a statistically
significant increase in point counts over time.
---------- Space Time Cube Characteristics -----------
Input feature time extent          1984-05-08 00:00:00
                                to 2018-10-13 00:00:00

Number of time steps                                 35
Time step interval                             1 year
Time step alignment                               End

First time step temporal bias                    51.64%
First time step interval                          after
                                   1983-11-01 00:00:00
                                       to on or before
                                   1984-11-01 00:00:00

Last time step temporal bias                      5.21%
Last time step interval                           after
                                   2017-11-01 00:00:00
                                       to on or before
                                   2018-11-01 00:00:00

Cube extent across space       (coordinates in meters)
Min X                                      419415.0000
Min Y                                     4426965.0000
Max X                                      446235.0000
Max Y                                     4468905.0000
Rows                                               233
Columns                                            149
Total bins                                     1215095
----------------------- COUNT ------------------------
Total number of locations                        34717
Locations with at least one point                14180
- associated bins                               496300
- % non-zero (sparseness)                         84.03
---- Summary Field - THERM_MEAN_SPATIAL_NEIGHBORS ----
% of locations excluded due to unfilled bins      7.52%
- Total number                                    2611
Total number of locations                        11569
- associated bins                               404915
% of included locations with estimated bins       8.71%
- Total number                                    1008
% of all bins that were estimated                3.68%
- Total number                                   14916
------------- Overall Data Trend - COUNT -------------
Trend direction                             Increasing
Trend statistic                                 4.3172
Trend p-value                                   0.0000
- Overall Data Trend - THERM_MEAN_SPATIAL_NEIGHBORS –
```

```
  Trend direction                              Increasing
  Trend statistic                                  2.7554
  Trend p-value                                    0.0059
  Completed script Create Space Time Cube By Aggregating Points...
```
Succeeded at Thursday, April 11, 2019 5:27:01 PM (Elapsed Time: 4 minutes 43 seconds)