

Remote Sensing and Cloud Computing to Support Lake Tahoe Nearshore Monitoring

Christopher Pearson¹
Justin L. Huntington^{1,2}

September 2019



¹Division of Hydrologic Sciences, Desert Research Institute

²Western Regional Climate Center, Desert Research Institute

Prepared for
Nevada Division of Environmental Protection



NEVADA DIVISION OF
**ENVIRONMENTAL
PROTECTION**

Cover Page Photo Credit: USGS, <https://www.usgs.gov/media/images/timeline-nearshore-periphyton-growth-lake-tahoe>

Executive Summary

Remote sensing of water quality in freshwater lakes and rivers continues to be a highly researched topic due to the great need and potential for monitoring over large areas and time periods. Results from previous studies suggest that remote sensing can be used to effectively monitor water quality in large water bodies such as Lake Tahoe. However, most previous studies have focused on mid-lake in-situ water quality measurements related to suspended constituents for developing correlations and making predictions with remote sensing data. Developing correlations and making predictions in the nearshore is potentially more challenging than at mid-lake locations due to reflectance from the nearshore lake bottom and the potential for mixed land and water reflectance for pixels near the water's edge. The objective of this study was to assess the potential for monitoring nearshore periphyton at Lake Tahoe, using Landsat satellite imagery that is atmospherically corrected using standard algorithms and is freely available for operational applications. In-situ periphyton data collected during both routine and synoptic sampling campaigns going as far back as 1984 was paired with cloud-free Landsat satellite imagery processed using the Google Earth Engine cloud computing platform in order to explore statistical relationships between Landsat surface reflectance and periphyton chlorophyll-a, with the goal of developing a predictive algorithm and enable the use of remote sensing for historical and operational monitoring of algae within the nearshore.

Results indicate that a universal algorithm for all seasons lacked sufficient correlation and skill in predicting water quality metrics associated with algae (e.g. plant biomass index and chlorophyll-a). Lack of correlation and predictive skill is due to a combination of variations in bottom characteristics and type, non-unique sources of reflectance, spatial heterogeneity at each in-situ sampling site, low signal to noise of surface reflectance from water, image geolocation accuracy, and image spatial resolution. However, site specific algorithms controlled for seasonality had acceptable correlation with in-situ periphyton chlorophyll-a concentration and showed promise for prediction. More specifically, application of season specific multivariate regressions with Landsat data demonstrated the potential for historical and operational prediction of water quality metrics during time periods with little to no in-situ measurements. Computed anomalies of satellite-based water quality metrics, both as maps and time series, also correlated well to periods of known high and low algae concentrations.

Major Findings

- Historical data analysis shows large seasonal variability with highest concentrations in the spring and winter
- The development of a universal algorithm that had acceptable correlation with in-situ periphyton chlorophyll-a concentrations was not possible due to variations in bottom type, spatial heterogeneity at each in-situ sampling site, low signal to noise of surface reflectance from water, image geolocation accuracy, and image spatial resolution
- Site and season specific algorithms using multiple Landsat surface reflectance bands and multivariate regression had higher correlation with in-situ periphyton chlorophyll-a concentrations

- Predictions of in-situ periphyton chlorophyll-a illustrate the potential for operational monitoring and gap filling of historic datasets, where the predictions have the same amount of temporal variability as the in-situ observations
- Relative periphyton chlorophyll concentration and biomass can be achieved through anomaly mapping and analysis of various Landsat water quality indices via the Climate Engine cloud computing application (app.climateengine.org)
- Historical and operational monitoring of nearshore periphyton chlorophyll-a and biomass may be possible at specific sites throughout the Tahoe Basin
- Targeted sampling during satellite overpass days, documenting GPS locations for each in-situ sample would likely improve site specific correlations and prediction accuracy
- Integrating the use of new free and operational satellite data, such as Sentinel 2 and Sentinel 3, with Landsat and into cloud computing platforms such as Climate Engine has the potential to greatly improve monitoring through improved temporal and spectral resolution

The use of remote sensing is a feasible option for monitoring nearshore water quality at Lake Tahoe, however, there are challenges. Integration and use of new satellite data combined with targeted sampling aligned with satellite image acquisition dates could yield long-term benefits for monitoring with satellite remote sensing. While other water quality remote sensing platforms are available, such as the use of piloted aircraft and drones, these platforms are expensive and extremely limited for long term operational monitoring and applications. We believe that investing in research and applications that combine in-situ data with freely available satellite remote sensing data is the best way for ensuring affordable, sustainable, and scalable long-term monitoring capabilities at Lake Tahoe.

Introduction

The waters that line the shores of Lake Tahoe (Figure 1) are highly valued for their aesthetic and recreational qualities. In recent times, Lake Tahoe's nearshore conditions with respect to algae growth have become more evident to visitors, residents, and resource management agencies. Possible degradation of nearshore quality through increased algal growth on rocks, reductions in water column transparency within mid-lake areas, and the establishment and spread of aquatic invasive species (AIS) threaten Tahoe as a waterbody with unique aesthetical and ecological functions and status, recreational destination, drinking water source, and asset to the local and regional economies. Understanding how to better manage Lake Tahoe's nearshore environment and the areas that affect it is a high priority to stakeholders and resource management agencies. An important part of improved management is improved knowledge of current status and trends. Developing and implementing cost-effective and reliable methods for assessing current status and long-term trends of nearshore water quality, and for assessing the effectiveness of new management strategies and actions is needed.

Through financial support provided by the Tahoe Science Program (TSP) and the Nearshore Agency Work Group (NAWG) comprised of the Nevada Division of Environmental Protection (NDEP), Lahontan Water Board, Tahoe Regional Planning Agency (TRPA), and United States Environmental Protection Agency (US EPA), a science team comprised of the Desert Research Institute (DRI), University of Nevada Reno (UNR) and the University of California Davis (UCD) collaborated with the TSP and the NAWG to synthesize available nearshore information and to develop a comprehensive strategy to monitor and assess nearshore health. The resulting Nearshore Evaluation and Monitoring Framework Report (Nearshore Report; Heyvaert et al. 2013) put forth a monitoring plan framework aimed at providing a comprehensive and integrated assessment of nearshore condition. However, monitoring the ten proposed metrics (turbidity, light transmission, chlorophyll, phytoplankton, periphyton, macrophytes, benthic macroinvertebrates, fish and crayfish, toxicity and harmful microorganisms) is estimated to cost on the order of \$500,000 annually. High costs are driven by the need for in-situ point monitoring techniques that are both time and labor intensive and limited in spatial and temporal extent – all common challenges with in-situ monitoring. Because nearshore conditions have the potential to be highly variable in space and time, in-situ field monitoring has limited ability to quantify this variability and shed insight on its cause.

Offering continuous coverage in time and space at reduced costs, remote sensing has been used to complement in-situ monitoring techniques and provide information for areas or time periods where no in-situ monitoring exists. Several studies have demonstrated the feasibility of estimating chlorophyll-a, colored dissolved organic matter (CDOM), and total suspended sediment concentrations (TSS) in lakes from multi-spectral satellite, airborne, and Unmanned Aerial System (UAS) platforms (Allan et al. 2011; Brezonik et al. 2005; Torbick et al. 2008; Zang et al. 2012). Steissberg et al. (2010) developed useful empirical relationships at Lake Tahoe between low spatial resolution satellite derived water quality indices and mid-lake in-situ water quality measurements of turbidity, Secchi disk, and chlorophyll-a. Additionally, previous studies by the PIs relating paired Landsat satellite imagery to in-situ water quality measurements for many western Nevada reservoirs have proven useful for evaluating spatial and temporal distributions and trends of TSS and Secchi disk (Pahl and Huntington, 2010).

Given the success of previous studies, assessing the potential for remote sensing to offer an effective, low-cost technique to monitor water quality over large areas and time periods, and to complement and support routine and synoptic in-situ measurements for Lake Tahoe's nearshore environment is warranted. Of specific interest is the ability to monitor Periphyton. Periphyton is form of attached algae that reflects both local nutrient loading and long-term environmental fluctuations (i.e. climate, temperature, runoff, etc.). Application of remote sensing techniques for monitoring water quality at Lake Tahoe is not without significant challenges, especially since Lake Tahoe has much greater clarity than lakes where remote sensing shown great promise for monitoring water quality. Challenges are associated with atmospheric correction and obtaining more signal than noise in measured at-surface reflectance, and the fact that the water column contains a mix of water quality constituents, such as colored dissolved organic matter (CDOM), chlorophyll-a, and TSS. More specifically, different combinations of CDOM, chlorophyll, and turbidity can return similar spectral signals, therefore inverse modeling is often required to simulate the optical properties of the water by varying constituent concentrations until the model results match the remotely sensed or measured optical properties. Empirical relationships can be made with historical in-situ water quality measurements and paired remote sensing data; however, many observations are needed covering a wide variety of water quality conditions.

Objectives

The objective of this study was to assess the potential for operational monitoring nearshore periphyton at Lake Tahoe using Landsat satellite imagery that is atmospherically corrected using standard algorithms and is freely available for operational applications. To do so in-situ periphyton data collected during both routine and synoptic sampling campaigns at 50 sites and at approximately monthly time steps, going as far back as 1984 was quality assured and controlled, organized, and paired with Landsat satellite imagery in order to explore statistical relationships between Landsat surface reflectance and periphyton chlorophyll-a, with the ultimate goal of developing a predictive algorithm to enable the use of remote sensing and new cloud computing technologies for historical and operational monitoring of periphyton within the nearshore.

Study Location

Lake Tahoe is a large freshwater lake located in the Sierra Nevada Mountains on the border of California and Nevada. The lake has a maximum depth of 501 m (1,645 ft) and surface area of 490 km² (191 mi²). Approximately two-thirds of Lake Tahoe Basin parent material is granitic and one-third is volcanic. The Lake Tahoe Watershed covers 1,310 km² (505 mi²) and drains into the lake via 63 tributaries. Runoff accounts for approximately half the annual water input, while direct precipitation to the lake in the form of rain and snow accounts for the other half. Lake Tahoe Basin climate is characterized by warm dry summers and cold wet winters. The majority of precipitation falls as snow between November and April with a distinct west to east rain shadow effect. Mean annual precipitation on the west side of the basin is approximately 86 cm, and almost double the 46 cm measured on the east side of the basin.

Lake Tahoe's large volume, surface area to watershed area ratio, and low nutrient soils make for extremely high purity water. However, increased nutrient and sediment loading have decreased clarity and increased algae growth since monitoring began in the 1960s. Secchi disk measurements in the center of the lake have decreased from 30.5 to 21.3 m, while trends in nearshore degradation and algae outbreaks are not well documented nor understood. The nearshore encompasses area from the lake's edge to a depth contour where the thermocline intersects the lake bed in mid-summer (minimum width of 350ft; Heyvaert, 2013). Tahoe's water quality and clarity support a unique ecosystem and a large tourism-based economy with the nearshore zone serving as the primary area for public interaction and visibility. Protecting Tahoe's environment and water quality has become public issue with multiple agencies and stakeholders supporting large efforts to track and mitigate further degradation.

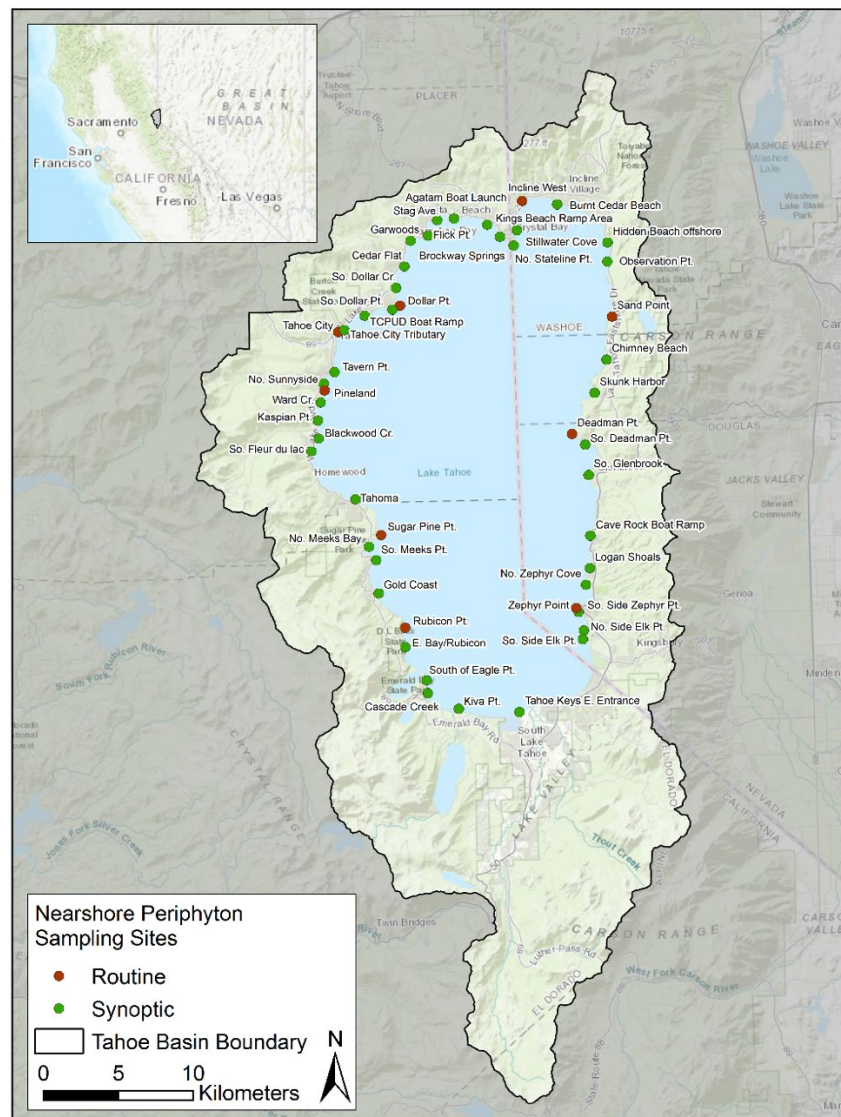


Figure 1: Map of the Lake Tahoe Basin showing both the Routine (red) and Synoptic (green) nearshore periphyton sampling locations.

Data

Nearshore water quality datasets collected and processed by UC Davis contained historical periphyton data from both Routine and Synoptic sampling campaigns. Throughout this report, chlorophyll-a refers to periphyton chlorophyll and is representative of algae biomass. Routine sampling occurs at nine nearshore sites throughout the lake on a monthly to quarterly basis, while spring synoptic sampling occurs at an additional forty locations during the period of the spring maximum (Figure 1). Sampling coverage varies from site to site, but extends as far back as 1982. All periphyton samples are collected at a consistent depth of 0.5 m regardless of lake level (TERC, 2017). The Periphyton dataset was processed for Quality Assurance and Control (QAQC) and organized for pairing with remote sensing datasets.

Landsat at-surface (SR) and top-of-atmosphere (TOA) data was extracted from the USGS Tier 1 collections via the Google Earth Engine Python Application Programming Interface (API) (Gorelick, 2017). The Google Earth Engine cloud computing platform provides a collections of historic satellite imagery and other geospatial datasets for near real-time access and on-the-fly processing. During the study, the USGS released the updated Tier 1 Landsat collections with improved data quality and calibrations (USGS, 2018a,b). Initial work and presentations were based on the older (now retired) collections; however, all results presented here are based on the updated collections.

Satellite band ratios and chlorophyll estimates were calculated from unique band combinations (detailed in methods section below). Table 1 shows a summary of total in-situ sample counts for each site, and the number of in-situ samples that were collected within 7, 5, and 2 days of a cloud-free satellite image for the Routine sampling sites. With increased temporal gaps between in-situ sample and satellite image acquisition dates the paired in-situ – satellite image sample counts increase, but the representativeness of the satellite data reduces. As detailed in the results section, regression analyses using the 2 day datasets consistently produced the highest correlations, making the use of 2 day datasets the most viable option for algorithm development.

Table 1: Summary of paired in-situ and cloud-free satellite image counts for satellite images acquired with 7, 5, and 2 days of in-situ sampling dates.

Station	Sample Type	Chlorophyll Sample Count	Valid Landsat Images w/in 7 days	Valid Landsat Images w/in 5 days	Valid Landsat Images w/in 2 days
Deadman Pt.	Routine	145	67	56	31
Dollar Pt.	Routine	153	72	58	35
Incline West	Routine	98	57	48	29
Pineland	Routine	160	73	58	33
Rubicon Pt.	Routine	151	69	54	28
Sand Point	Routine	145	66	55	32
Sugar Pine Pt.	Routine	89	52	38	20
Tahoe City	Routine	100	59	46	29
Zephyr Point	Routine	153	66	55	33

Methods

Historical Landsat surface reflectance images from 1984 to present for Path 43 and Row 33 were acquired and processed using the Google Earth Engine API and custom Python software developed as part of this project. The USGS converts Landsat TM (Landsat 5), ETM+ (Landsat 7), and OLI (Landsat 8) top-of-atmosphere (TOA) reflectance to at-surface reflectance (SR) using the Landsat ecosystem disturbance adaptive processing system (LEDAPS) (for TM and ETM+) and Landsat Surface Reflectance Code (LaSRC) (for OLI) atmospheric correction algorithms (Schmidt et al., 2013; USGSb, 2018). After atmospheric correction, Landsat TOA and SR images were visually (QAQCed) and sorted into cloudy and non-cloudy groupings in order to identify the highest quality images suitable for analysis. Cloud masks were also applied during the data processing using F-Mask (Zhu and Woodcock, 2012) to further ensure image quality.

Select band ratios commonly used as water quality indices were computed and included in addition to the individual band reflectance. Previous water quality studies have identified good statistical relationships between satellite band ratios and both chlorophyll-and turbidity (Allen, 2011; Barrett, 2016; Kloiber et. al., 2002). The most commonly used ratio for chlorophyll mapping is the Green to Blue ratio, however, different ratios can also be of value depending on the season and target constituent. For example, NASA's Ocean Chlorophyll (OC) algorithm suite relies on empirical relationships using Blue to Green ratio (NASA, 2014). Chlorophyll strongly absorbs in the red and blue bands and reflects in the green and near infrared (NIR) bands (Barrett, 2016). Historically, blue reflectance has been associated with healthy systems, red reflectance has been associated with high sediment/turbid systems, and green reflectance has been associated with algal prone systems (Waxter, 2014). Ratios including the NIR band have also shown good skill in estimation of chlorophyll in turbid waters and coastal areas (Mohamed, 2015).

Landsat band data and band ratios and the NASA's Ocean Chlorophyll 2 (OC2) predicted chlorophyll were computed, spatially averaged (based on a weighted reducer in Google Earth Engine) and extracted for each image/sample pair using a 30m buffer around the sampling coordinates. Earth Engine's weighted reducer calculates an area-weighted average of all pixels falling within the 30m buffered zone. Bands processed include the Blue (450-510 nm), Green (530-590 nm), Red (640-670 nm), Near Infrared (NIR; 850-880 nm), Shortwave Infrared 1 (SWIR1; 1570-1650 nm), and Shortwave Infrared 2 (SWIR2; 2110-2290 nm) (note: wavelengths apply to the Landsat 8 platform). Both SR and TOA reflectance data were processed in order to assess the utility and effectiveness of operational atmospheric correction algorithms. Each in-situ water quality sample was paired with the closest in time cloud-free Landsat image.

After satellite and in-situ data were processed and organized, statistical analyses were performed to investigate correlations between the datasets. All statistical analyses were performed using custom Python scripts, including multi-variate, linear, and non-parametric regression (i.e. Spearman's Rho), outlier analyses (1.5 Interquartile Rule), and trend tests (Kendall's Tau). Specific python packages include pandas for data management (McKinney, 2010), SciPy (Oliphant, 2007; Millman and Aivasis, 2011) and sklearn (Pedregosa et. al., 2011) for regression and non-parametric trend testing, and Matplotlib for data graphing and visualization (Hunter, 2007). The following band ratios were included in statistical analyses -

Green to Blue (GtoB), Red to Blue (RtoB), Red to Green (RtoG), Green to NIR (GtoN), Red to NIR (RtoN), Blue to NIR (BtoN), and OC2 chlorophyll. The GtoB ratio has shown good agreement with algae concentrations collected from ocean and inland waters and is the basis for Ocean Chlorophyll algorithms developed by NASA (NASA, 2014).

Results and Discussion

The following section highlights results of in-situ samples, statistical correlations between in-situ and Landsat water quality metrics, the development of site-specific algorithms, and application of those algorithms for predicting periphyton chlorophyll-a concentrations. Nearshore in-situ periphyton chlorophyll-a concentrations show a distinct seasonality at all sites, with highest concentrations in winter and spring and lower concentrations in summer and fall (Figure 2). Minimum and maximum chlorophyll concentrations varied spatially, with higher concentration sites having increased seasonal variability.

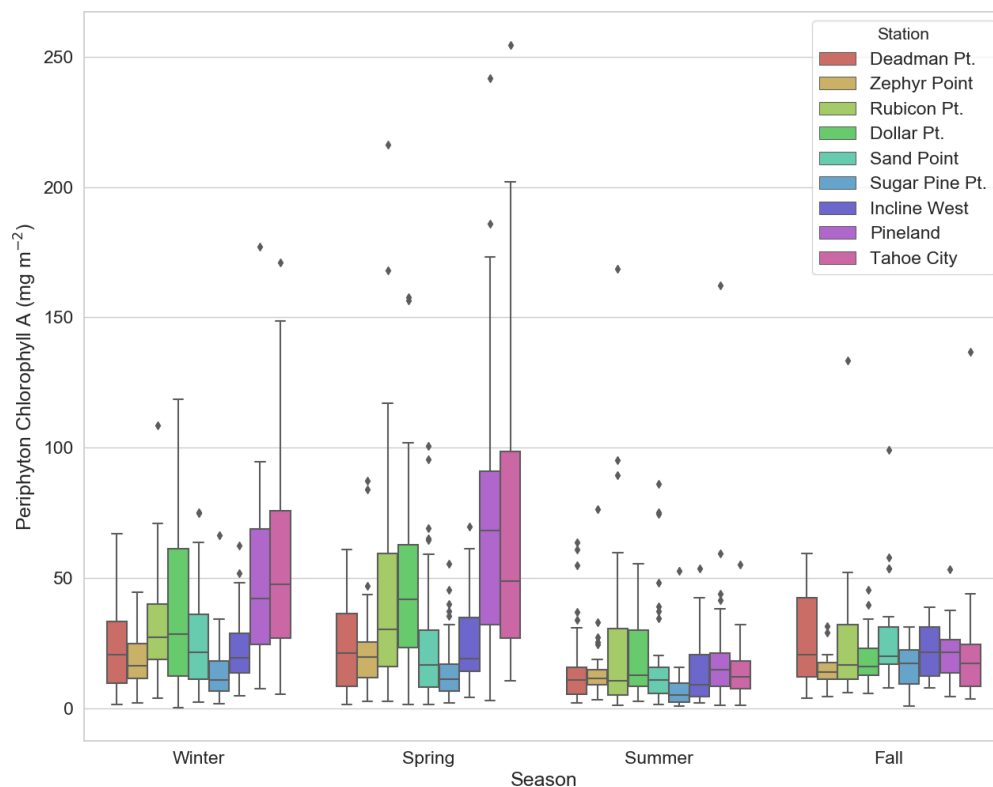


Figure 2: Boxplots of in-situ seasonal periphyton chlorophyll-a concentrations for each Routine sampling site. The center line within each box represents the median of the dataset, while each box extends to the quartiles of the datasets. Whiskers extend to 1.5 times the interquartile range with outliers represented as points falling outside.

Both in-situ periphyton chlorophyll-a and satellite band data were log-normally distributed and were therefore log-transformed prior to all statistical analyses. Linear regression

(i.e. Pearson's r-value) was used to assess correlation between periphyton chlorophyll-a and each satellite band/ratio. In general, SR datasets showed higher correlations and overall tighter relationships with nearshore water quality data over the TOA dataset. TOA reflectance is impacted by backscatter and absorption related to particulates, aerosols, and water vapor throughout atmospheric profile. SR datasets are processed to remove these impacts and provide a more accurate and temporally consistent representation of the surface reflectance. The improved SR correlations validate the effectiveness of the atmospheric correction algorithms, however, neither SR or TOA datasets showed sufficient statistical strength capable for developing a universal (i.e. lake-wide) algorithm for predicting periphyton chlorophyll-a. For example, Figure 3 shows large scatter in chlorophyll-a concentrations relative to the GtoB ratio when including data from all sites and all time periods for Landsat images acquired within 2-days of a cloud free satellite image. Site-specific physical characteristics such as variability in bottom-type and water depth, and differences in the seasonality of periphyton chlorophyll-a concentrations are likely significant contributors to the overall scatter illustrated in Figure 3.

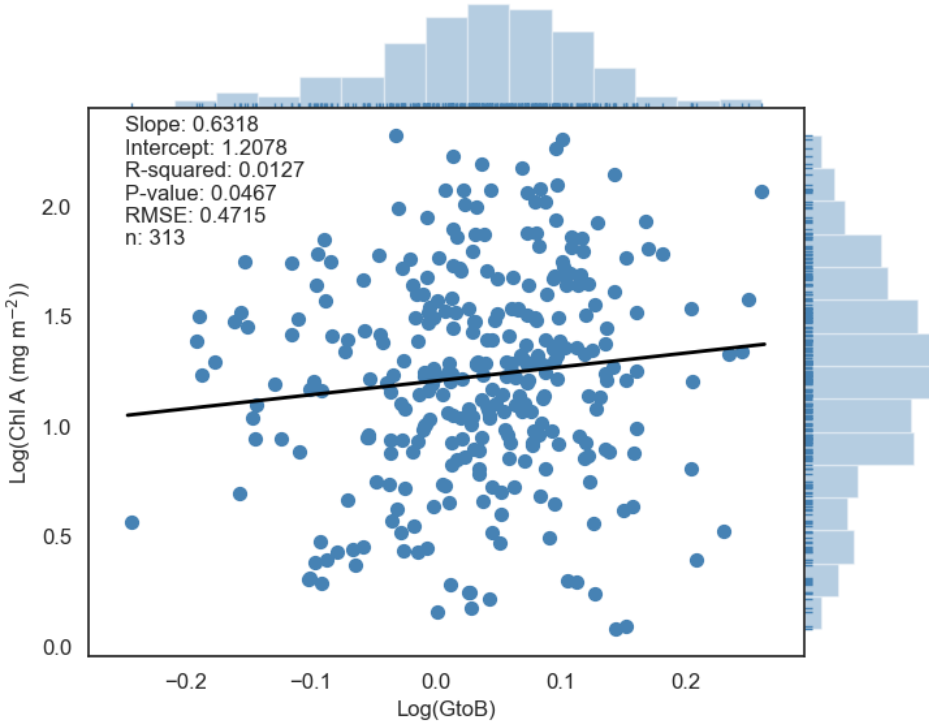


Figure 3: Regression analysis of the Green to Blue band ratio (GtoB) versus periphyton chlorophyll-a concentrations (Chl A) from all routine sites where samples were collected within 2-days of a cloud free image. Although correlated ($p\text{-value} < 0.05$), the relationship is not strong enough to support predictive modeling across all sites.

In-situ periphyton sampling at a consistent depth of 0.5m means that spatially coincident satellite pixels include reflectance not only from periphyton biomass, but also from the lake bottom. Intra-site variability in bottom substrate type, water depth, and heterogeneity reduces statistical power when assessing all sites together. To reduce the impact of intra-site variability, site-specific statistical analyses were performed and resulted in marked improvements,

highlighting the potential for site-specific algorithms to be developed at the annual time scale (Table 2).

Table 2: Site specific linear regression analysis results between log-transformed in-situ periphyton chlorophyll-a (Chl A) and Landsat SR bands and band ratios using all in-situ samples collected within 2-days of a cloud free image acquisition. Top) results with abs(r-values) > 0.5 are highlighted in green and results with abs(r-values) > 0.4 are shown in red. Bottom) results with p-values < 0.05 are highlighted in green and results with p-values < 0.1 are shown in red.

r-values	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point
n	31	35	29	33	28	32	20	29	33
Blue	-0.394	0.007	0.169	0.230	0.141	-0.446	0.187	0.202	0.255
Green	-0.377	0.054	0.261	0.318	0.155	-0.358	0.264	0.158	0.223
Red	-0.322	0.114	0.361	0.319	0.177	-0.380	0.336	0.162	0.234
NIR	-0.334	0.032	0.273	0.134	0.145	-0.255	0.374	0.033	0.048
SWIR1	-0.180	0.093	0.643	-0.023	0.133	-0.026	0.363	-0.127	0.098
SWIR2	-0.233	0.084	0.697	0.050	-0.003	-0.151	0.294	-0.072	0.183
Green/Blue	-0.075	0.174	0.248	0.127	0.044	0.181	0.267	-0.463	-0.093
Red/Blue	-0.253	0.220	0.611	0.318	0.168	-0.134	0.405	-0.251	0.053
Green/Red	0.248	-0.177	-0.656	-0.266	-0.181	0.331	-0.358	-0.072	-0.203
Blue/NIR	0.244	-0.045	-0.068	-0.059	-0.013	-0.043	-0.411	0.430	0.146
Green/NIR	0.225	0.007	0.088	-0.028	0.009	0.054	-0.378	0.368	0.169
Red/NIR	-0.176	0.210	0.433	0.105	0.164	-0.366	-0.341	0.441	0.384
Blue/SWIR1	-0.046	-0.078	-0.072	0.142	0.039	-0.105	-0.283	0.197	-0.009
Green/SWIR1	-0.073	-0.041	-0.012	0.188	0.049	-0.088	-0.273	0.169	-0.031
Red/SWIR1	-0.207	0.039	0.132	0.289	0.090	-0.180	-0.102	0.179	0.011

p-values	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point
n	31	35	29	33	28	32	20	29	33
Blue	0.028	0.969	0.382	0.257	0.475	0.011	0.431	0.294	0.153
Green	0.036	0.758	0.171	0.113	0.431	0.044	0.260	0.414	0.213
Red	0.077	0.516	0.054	0.112	0.369	0.032	0.148	0.402	0.190
NIR	0.066	0.854	0.151	0.513	0.462	0.158	0.104	0.864	0.792
SWIR1	0.333	0.595	0.000	0.910	0.501	0.887	0.116	0.511	0.587
SWIR2	0.207	0.632	0.000	0.807	0.988	0.410	0.209	0.709	0.307
Green/Blue	0.688	0.318	0.195	0.537	0.823	0.322	0.256	0.011	0.607
Red/Blue	0.170	0.204	0.000	0.113	0.393	0.465	0.076	0.190	0.770
Green/Red	0.179	0.309	0.000	0.190	0.356	0.064	0.121	0.711	0.258
Blue/NIR	0.186	0.797	0.727	0.776	0.948	0.815	0.072	0.020	0.417
Green/NIR	0.225	0.968	0.649	0.891	0.963	0.767	0.100	0.050	0.346
Red/NIR	0.344	0.227	0.019	0.611	0.405	0.039	0.141	0.017	0.027
Blue/SWIR1	0.804	0.658	0.712	0.488	0.844	0.568	0.227	0.305	0.963
Green/SWIR1	0.698	0.817	0.951	0.359	0.803	0.634	0.243	0.381	0.864
Red/SWIR1	0.263	0.822	0.494	0.153	0.649	0.324	0.669	0.354	0.952

In order to further refine correlations beyond site-specific analyses, the data was further subset and analyzed by season. Regression analyses were repeated for all sites and band/ratio

combinations for winter (Dec-Feb), spring (Mar-May), summer (Jun-Aug), and fall (Sept-Nov). Site/season combinations with less than 5 samples were not assessed. Seasonal site-specific regression analyses showed additional improvements with multiple site/season combinations showing potential for predictive algorithm development (Table 3).

Table 3: Site specific linear regression analysis results between log-transformed in-situ periphyton chlorophyll-a (Chl A) and Landsat SR bands and band ratios using all in-situ samples collected within 2-days of a cloud free image acquisition separated by season. Top) Results with abs(r-values) > 0.5 are highlighted in green and results with abs(r-values) > 0.4 are shown in red. Bottom) Results with p-values < 0.05 are highlighted in green and results with p-values < 0.1 are shown in red.

Winter (Dec-Feb)										Summer (Jun-Aug)									
r-values	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point	r-values	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point
n	8	11	8	8	11	10	9	9	10	n	9	7	9	8	7	10	3	5	10
Blue	-0.264	-0.205	-0.643	0.007	0.129	-0.202	0.223	-0.439	-0.177	Blue	-0.650	-0.216	-0.280	-0.097	-0.307	-0.136		0.467	0.286
Green	-0.340	-0.164	-0.602	0.176	0.054	-0.158	0.342	-0.454	-0.022	Green	-0.656	0.475	0.647	-0.015	-0.404	0.203		0.450	0.370
Red	-0.281	-0.211	-0.584	0.133	0.069	-0.402	0.317	-0.444	0.020	Red	-0.632	0.451	0.629	0.197	-0.298	0.111		0.428	0.435
NIR	-0.293	-0.258	-0.563	0.220	-0.168	-0.128	0.325	-0.557	-0.074	NIR	-0.683	0.254	0.490	-0.172	-0.428	0.168		0.305	0.459
SWIR1	0.128	-0.067	0.712	-0.318	-0.428	0.010	0.661	0.128	0.175	SWIR1	-0.215	-0.137	0.417	0.362	0.550	0.386		0.169	0.509
SWIR2	-0.121	0.142	0.766	-0.270	-0.222	-0.262	0.747	0.238	0.588	SWIR2	-0.190	-0.063	0.611	0.345	0.382	0.191		0.266	0.501
Green/Blue	-0.252	0.274	0.690	0.397	-0.519	0.142	0.606	0.201	0.281	Green/Blue	-0.073	0.823	0.882	0.233	-0.122	0.394		0.057	0.116
Red/Blue	-0.276	-0.122	0.581	0.307	-0.313	-0.252	0.429	0.273	0.262	Red/Blue	-0.573	0.676	0.855	0.382	-0.222	0.246		0.204	0.372
Green/Red	0.239	0.437	-0.008	-0.073	-0.130	0.511	-0.237	-0.268	-0.132	Green/Red	0.559	-0.403	-0.539	-0.338	0.177	0.049		-0.291	-0.516
Blue/NIR	0.300	0.226	-0.679	-0.409	0.660	-0.010	-0.375	-0.122	-0.089	Blue/NIR	0.601	-0.470	-0.601	0.190	0.337	-0.308		-0.006	-0.277
Green/NIR	0.224	0.322	-0.587	-0.219	0.544	0.078	-0.255	-0.092	0.064	Green/NIR	0.577	0.183	0.302	0.204	0.412	-0.025		0.017	-0.364
Red/NIR	-0.213	0.069	-0.535	-0.279	0.537	-0.399	-0.226	0.007	0.119	Red/NIR	-0.198	0.531	0.525	0.395	0.090	-0.116		0.174	0.056
Blue/SWIR1	-0.268	-0.159	-0.773	0.229	0.282	-0.055	-0.422	-0.327	-0.178	Blue/SWIR1	-0.199	0.097	-0.644	-0.527	-0.657	-0.427		0.241	-0.364
Green/SWIR1	-0.367	-0.117	-0.757	0.352	0.231	-0.038	-0.388	-0.332	-0.147	Green/SWIR1	-0.219	0.349	0.652	-0.533	-0.767	-0.368		0.299	-0.489
Red/SWIR1	-0.310	-0.170	-0.744	0.307	0.237	-0.153	-0.273	-0.325	-0.137	Red/SWIR1	-0.636	0.599	0.729	-0.433	-0.630	-0.399		0.551	-0.299
Spring (Mar-May)										Fall (Sept-Nov)									
r-values	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point	r-values	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point
n	10	13	9	7	8	9	7	12	9	n	5	6	3	5	4	4	2	4	5
Blue	-0.395	0.113	0.294	0.018	0.482	-0.630	0.528	0.342	0.742	Blue	-0.413	-0.111		0.823					0.691
Green	-0.405	0.069	0.534	0.196	0.759	-0.576	0.576	0.350	0.652	Green	0.293	0.336		0.965					0.746
Red	-0.318	0.219	0.771	0.232	0.827	-0.579	0.764	0.373	0.669	Red	-0.210	0.582		0.938					0.669
NIR	-0.367	0.228	0.908	0.049	0.781	-0.594	0.874	0.334	0.419	NIR	-0.094	0.839		0.713					0.275
SWIR1	-0.488	0.380	0.892	-0.272	0.841	-0.554	0.399	-0.016	0.187	SWIR1	0.414	0.162		0.784					0.682
SWIR2	-0.358	0.281	0.899	-0.147	0.678	-0.352	0.178	-0.023	0.200	SWIR2	-0.363	0.071		0.771					0.687
Green/Blue	-0.053	-0.114	0.245	0.314	0.244	0.436	0.117	-0.263	-0.556	Green/Blue	0.720	0.475		0.666					0.759
Red/Blue	-0.223	0.260	0.853	0.366	0.624	-0.448	0.773	0.049	-0.268	Red/Blue	-0.093	0.735		0.877					0.654
Green/Red	0.213	-0.323	-0.926	-0.286	-0.873	0.519	-0.816	-0.289	-0.569	Green/Red	0.367	-0.765		-0.613					-0.539
Blue/NIR	0.276	-0.252	-0.374	-0.076	-0.447	0.341	-0.905	0.256	0.613	Blue/NIR	-0.131	-0.673		-0.398					0.269
Green/NIR	0.261	-0.284	-0.385	0.060	-0.496	0.417	-0.902	0.221	0.691	Green/NIR	0.221	-0.734		-0.223					0.654
Red/NIR	-0.038	-0.053	0.485	0.197	0.393	-0.352	-0.862	0.452	0.841	Red/NIR	-0.207	-0.334		0.041					0.923
Blue/SWIR1	0.066	-0.331	-0.347	0.264	-0.455	-0.386	-0.327	0.230	0.441	Blue/SWIR1	-0.579	-0.253		-0.604					-0.679
Green/SWIR1	0.059	-0.344	-0.365	0.409	-0.553	-0.318	-0.339	0.225	0.338	Green/SWIR1	-0.192	-0.049		-0.587					-0.624
Red/SWIR1	-0.053	-0.120	0.159	0.449	-0.061	-0.422	0.168	0.254	0.508	Red/SWIR1	-0.374	0.188		-0.468					-0.695

	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point
p-values									
n	8	11	8	8	11	10	9	9	10
Blue	0.528	0.545	0.085	0.987	0.705	0.576	0.564	0.237	0.625
Green	0.409	0.630	0.115	0.677	0.876	0.663	0.368	0.220	0.951
Red	0.501	0.533	0.128	0.754	0.840	0.249	0.405	0.232	0.956
NIR	0.482	0.444	0.146	0.601	0.621	0.724	0.394	0.120	0.840
SWIR1	0.762	0.845	0.048	0.442	0.189	0.978	0.053	0.742	0.629
SWIR2	0.775	0.678	0.027	0.517	0.512	0.464	0.021	0.537	0.074
Green/Blue	0.547	0.415	0.059	0.330	0.102	0.696	0.084	0.604	0.432
Red/Blue	0.508	0.721	0.131	0.459	0.349	0.482	0.250	0.478	0.464
Green/Red	0.569	0.179	0.984	0.864	0.703	0.131	0.539	0.486	0.716
Blue/NIR	0.471	0.503	0.064	0.315	0.027	0.979	0.320	0.754	0.807
Green/NIR	0.593	0.335	0.127	0.602	0.083	0.831	0.508	0.813	0.861
Red/NIR	0.612	0.841	0.172	0.503	0.088	0.254	0.558	0.986	0.743
Blue/SWIR1	0.521	0.641	0.025	0.585	0.401	0.879	0.258	0.391	0.624
Green/SWIR1	0.371	0.733	0.030	0.392	0.495	0.917	0.303	0.383	0.686
Red/SWIR1	0.455	0.617	0.034	0.459	0.483	0.674	0.477	0.393	0.706

Spring (Mar-May)

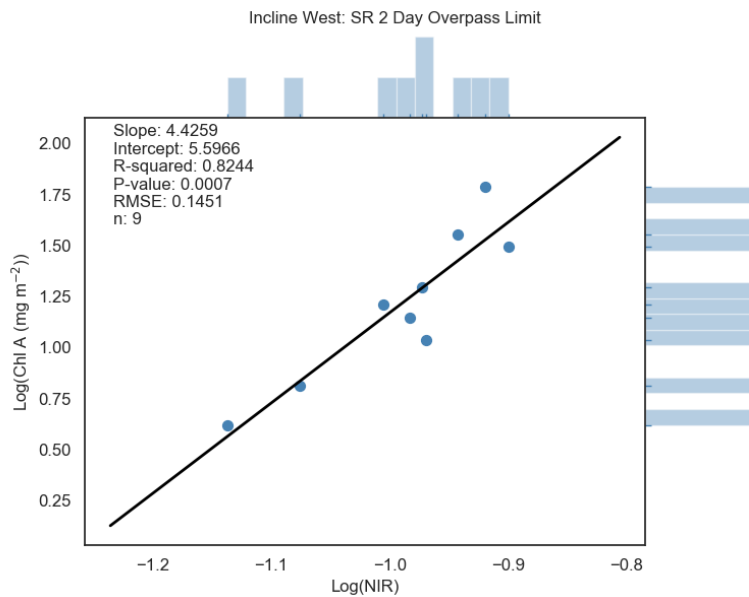
	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point
p-values									
n	10	13	9	7	8	9	7	12	9
Blue	0.259	0.713	0.443	0.969	0.227	0.069	0.223	0.277	0.022
Green	0.245	0.823	0.139	0.673	0.029	0.105	0.176	0.265	0.057
Red	0.370	0.471	0.015	0.617	0.011	0.102	0.045	0.233	0.049
NIR	0.296	0.454	0.001	0.917	0.022	0.092	0.010	0.289	0.262
SWIR1	0.153	0.200	0.001	0.555	0.009	0.121	0.375	0.960	0.629
SWIR2	0.310	0.353	0.001	0.752	0.065	0.353	0.703	0.944	0.607
Green/Blue	0.885	0.711	0.525	0.493	0.561	0.241	0.804	0.408	0.120
Red/Blue	0.536	0.392	0.004	0.420	0.098	0.226	0.041	0.881	0.486
Green/Red	0.554	0.281	0.000	0.534	0.005	0.152	0.025	0.363	0.110
Blue/NIR	0.441	0.406	0.321	0.871	0.267	0.369	0.005	0.423	0.079
Green/NIR	0.466	0.347	0.306	0.899	0.211	0.264	0.006	0.490	0.039
Red/NIR	0.916	0.862	0.186	0.673	0.336	0.353	0.013	0.140	0.005
Blue/SWIR1	0.856	0.269	0.360	0.567	0.257	0.305	0.474	0.472	0.235
Green/SWIR1	0.872	0.250	0.334	0.363	0.155	0.404	0.457	0.482	0.373
Red/SWIR1	0.884	0.697	0.683	0.313	0.887	0.258	0.718	0.426	0.162

	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point
p-values									
n	9	7	9	8	7	10	3	5	10
Blue	0.058	0.642	0.465	0.819	0.502	0.7086		0.427	0.424
Green	0.055	0.282	0.060	0.971	0.369	0.5738		0.447	0.292
Red	0.068	0.310	0.069	0.639	0.516	0.7595		0.473	0.208
NIR	0.043	0.582	0.181	0.684	0.338	0.6423		0.618	0.182
SWIR1	0.579	0.769	0.264	0.379	0.201	0.2706		0.786	0.133
SWIR2	0.625	0.894	0.080	0.403	0.397	0.598		0.666	0.140
Green/Blue	0.851	0.023	0.002	0.578	0.794	0.2598		0.928	0.749
Red/Blue	0.107	0.095	0.003	0.351	0.633	0.4941		0.742	0.290
Green/Red	0.118	0.370	0.134	0.413	0.705	0.8925		0.635	0.127
Blue/NIR	0.087	0.287	0.087	0.653	0.460	0.3863		0.993	0.438
Green/NIR	0.104	0.694	0.429	0.628	0.358	0.9459		0.978	0.301
Red/NIR	0.609	0.220	0.147	0.333	0.848	0.7498		0.780	0.879
Blue/SWIR1	0.609	0.835	0.061	0.180	0.109	0.219		0.696	0.300
Green/SWIR1	0.572	0.443	0.057	0.174	0.044	0.2961		0.625	0.151
Red/SWIR1	0.066	0.155	0.026	0.284	0.130	0.2534		0.336	0.402

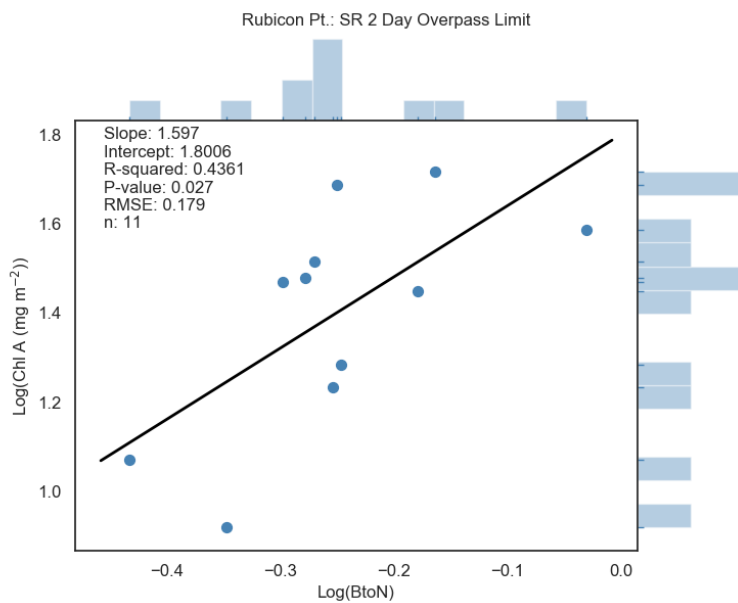
Fall (Sept-Nov)

	Deadman Pt.	Dollar Pt.	Incline West	Pineland	Rubicon Pt.	Sand Point	Sugar Pine Pt.	Tahoe City	Zephyr Point
p-values									
n	5	6	3	5	4	4	2	4	5
Blue	0.489	0.834		0.087					0.196
Green	0.633	0.515		0.008					0.148
Red	0.735	0.225		0.018					0.217
NIR	0.880	0.037		0.177					0.654
SWIR1	0.488	0.759		0.117					0.204
SWIR2	0.548	0.894		0.127					0.200
Green/Blue	0.171	0.341		0.220					0.136
Red/Blue	0.882	0.096		0.051					0.231
Green/Red	0.544	0.076		0.272					0.348
Blue/NIR	0.834	0.143		0.507					0.661
Green/NIR	0.721	0.096		0.719					0.231
Red/NIR	0.738	0.518		0.947					0.025
Blue/SWIR1	0.306	0.629		0.281					0.207
Green/SWIR1	0.758	0.927		0.298					0.260
Red/SWIR1	0.535	0.721		0.427					0.193

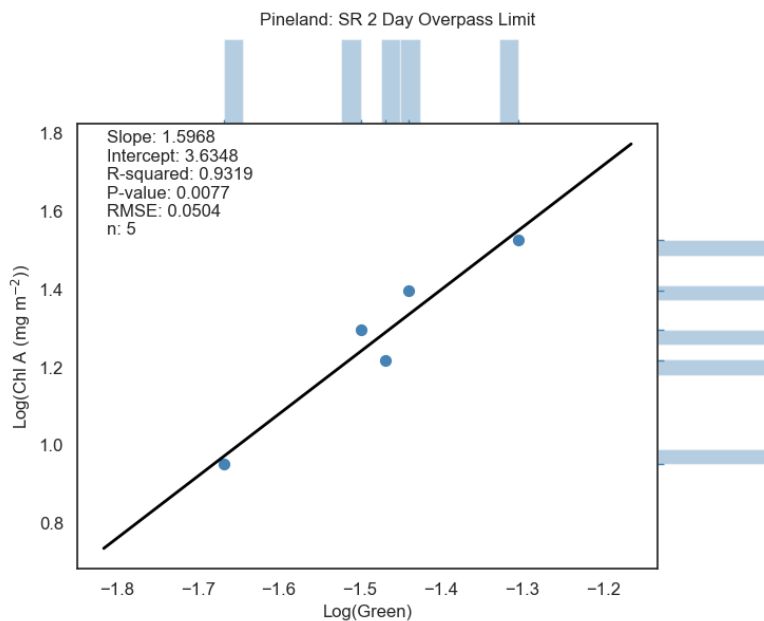
In general, spring resulted in the highest correlations, likely due to higher concentrations of chlorophyll-a providing increased reflectance well beyond the noise. While winter also showed relatively high concentrations, correlations were hindered by the limited number of cloud-free satellite images available during winter. Figure 4 highlights a subset of site and season-specific correlations between in-situ and Landsat SR reflectance bands and band ratios.



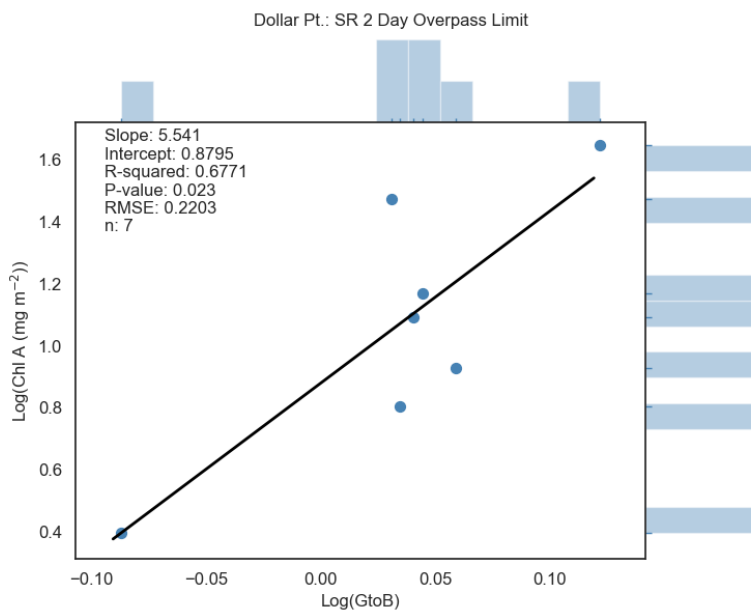
Incline West: Spring, NIR



Rubicon: Winter, Blue to NIR



Pineland: Fall, Green



Dollar Point: Summer, GtoB

Figure 4: Regression plots and summary statistics for a select group of station/season combinations. Each site and season combination correlate with different bands and band ratios making the development of a universal, single basin-wide algorithm not feasible.

Sub-setting site-specific imagery and in-situ data into monthly time periods helps to reduce potential effects from atmospheric conditions (i.e variability in particulates, aerosols, and water vapor within the atmosphere), turbidity, and periphyton bleaching, however, it also reduces the number of viable samples collected within 2 days of a cloud-free satellite image. This

threshold was relaxed to include imagery acquired within 5- and 7-days of in-situ samples, however, while there were more data points, correlations were reduced with higher temporal separation. Further complicating the analysis is the fact that different combinations of chlorophyll and turbidity can yield similar reflectance. Figure 5 illustrates how varying concentrations of chlorophyll-a and turbidity can result in non-unique spectral signals (i.e. colors).

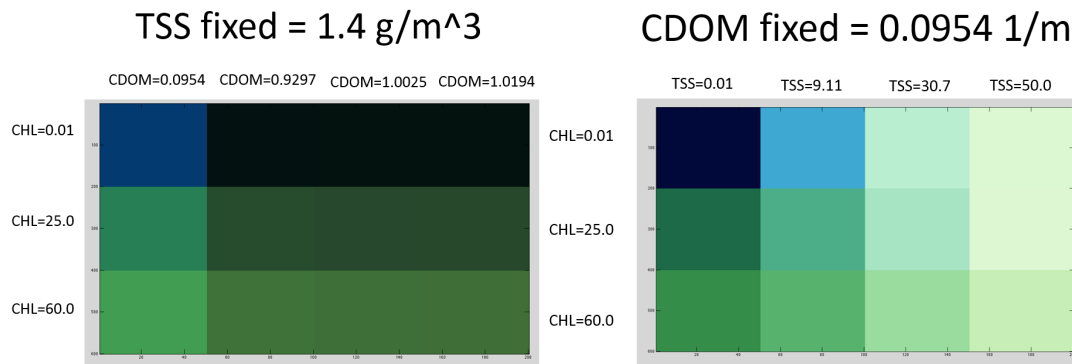


Figure 5: Diagram demonstrating the non-unique nature of chlorophyll/turbidity optics. Different combinations of colored dissolved organic matter (CDOM), chlorophyll, and turbidity can return similar spectral signals or colors (figure modified from Concha, 2015).

In addition to the non-unique nature of hydrologic optics, previous studies have noted fluctuations in chlorophyll content due to solar bleaching effects (Ichimura, 1959). Specifically, Glooschenko and Blanton (1977) noted a dependence on light intensity with higher solar input causing more pronounced fluctuations during a chlorophyll study in Lake Ontario. Previous studies at Lake Tahoe note that nearshore periphyton changes in state and condition, including both green and slimy/smooth states, and white and carpet-like states (State of the Lake Report, 2019). Periphyton bleaching at Lake Tahoe is potentially higher than other areas such as Lake Ontario, due to its high elevation and relatively clear waters, allowing for more shortwave radiation to penetrate the water column.

While site-specific and seasonal correlations of in-situ periphyton chlorophyll-a and Landsat reflectance data showed promise for predicting periphyton chlorophyll-a, ideally statistical power of the predictive model would be improved before used for operational monitoring. To explore improving the statistical power of site specific models, multivariate regression was applied using multiple bands and band ratios at select site and season combinations. Multivariate analysis was performed using the three bands with the highest individual correlations to periphyton chlorophyll-a for specific sites and timeframes. In general, multivariate regression based on multiple SR bands increased correlations improving overall predictive skill, however, rigorous optimization to find what site-band-season combinations yield the best skill is required. This rigorous optimization along with inclusion of new satellite data (i.e. Sentinel 2 and 3) should be explored as part of follow-on work to this study. An example result of the multivariate analysis for Incline West is shown in Table 4, where a single algorithm based on multivariate regression between SWIR2, RtoN, and RtoB and in-situ

periphyton data (i.e. including all data all seasons) was assessed (Table 4; adjusted r-squared: 0.574, p-value: <0.01) and then applied to evaluate Landsat predictions against the in-situ data.

Table 4: Incline West multivariate regression results between all available periphyton chlorophyll-a data and the corresponding Landsat SWIR2, RtoN, and RtoB bands and band ratios at the Incline West sampling site.

OLS Regression Results						
Dep. Variable:	Chl_A	R-squared:	0.620			
Model:	OLS	Adj. R-squared:	0.574			
Method:	Least Squares	F-statistic:	13.58			
Date:	Tue, 03 Sep 2019	Prob (F-statistic):	1.87e-05			
Time:	13:17:08	Log-Likelihood:	2.8604			
No. Observations:	29	AIC:	2.279			
Df Residuals:	25	BIC:	7.748			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	2.9262	0.780	3.750	0.001	1.319	4.533
SWIR2	1.2358	0.610	2.027	0.053	-0.020	2.492
RtoN	1.4477	0.591	2.449	0.022	0.230	2.665
RtoB	1.9510	0.874	2.233	0.035	0.152	3.750
Omnibus:	4.922	Durbin-Watson:	0.810			
Prob(Omnibus):	0.085	Jarque-Bera (JB):	3.352			
Skew:	-0.786	Prob(JB):	0.187			
Kurtosis:	3.548	Cond. No.	46.2			

Application of the Incline West multivariate regression model is illustrated in Figure 6. Comparing the time series of modeled versus measured in-situ chlorophyll-a concentrations shows relatively good agreement for time periods with overlapping data (Figure 6; Top). Landsat estimated chlorophyll-a concentrations follow the general five to ten-year variability of in-situ chlorophyll-a. Additionally, the Landsat estimated chlorophyll-a has the same level of temporal variability as the in-situ data at short time scales. These results suggest that the Landsat estimated chlorophyll-a can be used to fill in periods where in-situ data is limited or missing. Computing quarterly means of the Landsat time series reveals clear periodicity in predicted chlorophyll-a concentrations that track the in-situ data well (Figure 6; Bottom).

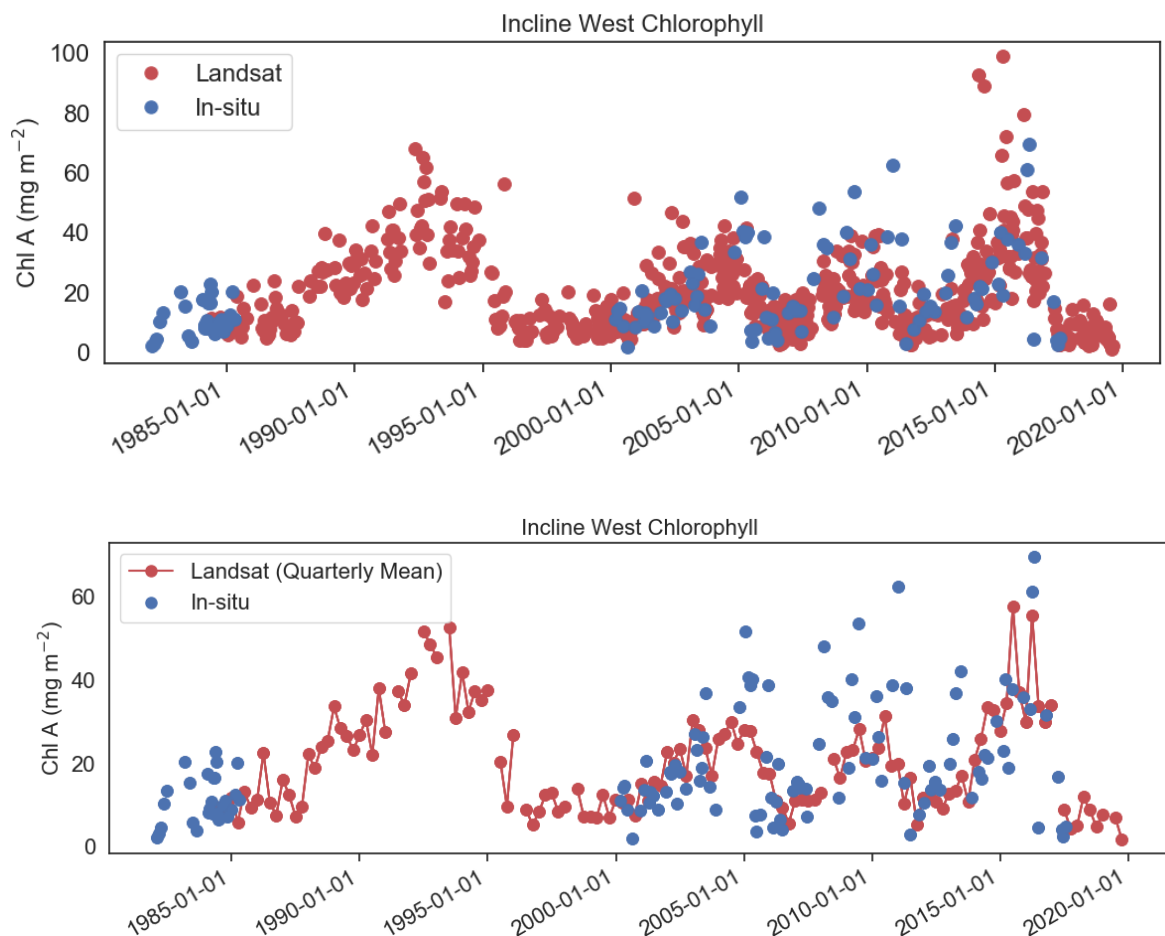


Figure 6: Top) Time series of in-situ (blue) and modeled (red) periphyton Chl A at the Incline West sampling site based on a multivariate regression using the SWIR2, RtoN, and RtoB bands and band ratios. Bottom) In-situ and modeled periphyton chlorophyll-a time series with Landsat estimates filtered to quarterly means.

Multivariate regression for site and season-specific combinations also shows potential for operational monitoring (Table 5: Rubicon Point Example). Similar to Incline West, a multivariate regression was performed for Rubicon Point, but developing separate models for each season, and applying those models to respective Landsat data and seasons to predict periphyton chlorophyll-a. Figure 7 illustrates a time series of Landsat estimated mean and maximum springtime (March-May) periphyton chlorophyll-a at the Rubicon Point site. Summarizing multiple predictions or in-situ samples over specific time periods (e.g. springtime) has the potential to reduce the temporal variability, and making interannual variability and possible trends more apparent.

Table 5: Multivariate regression results between spring (Mar, Apr, May) periphyton Chl A data and the corresponding Landsat SWIR1, Red, and Green bands and band ratios at the Rubicon Point sampling site..

OLS Regression Results						
Dep. Variable:	chl_A	R-squared:	0.810			
Model:	OLS	Adj. R-squared:	0.668			
Method:	Least Squares	F-statistic:	5.686			
Date:	Fri, 06 Sep 2019	Prob (F-statistic):	0.0632			
Time:	09:19:25	Log-Likelihood:	0.98207			
No. Observations:	8	AIC:	6.036			
Df Residuals:	4	BIC:	6.354			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	4.1143	2.826	1.456	0.219	-3.731	11.959
SWIR1	1.5946	1.628	0.980	0.383	-2.924	6.114
Red	6.9644	5.149	1.352	0.248	-7.333	21.261
Green	-7.4741	6.835	-1.094	0.336	-26.451	11.503
Omnibus:	1.245	Durbin-Watson:	1.866			
Prob(Omnibus):	0.537	Jarque-Bera (JB):	0.726			
Skew:	-0.354	Prob(JB):	0.695			
Kurtosis:	1.705	Cond. No.	224.			

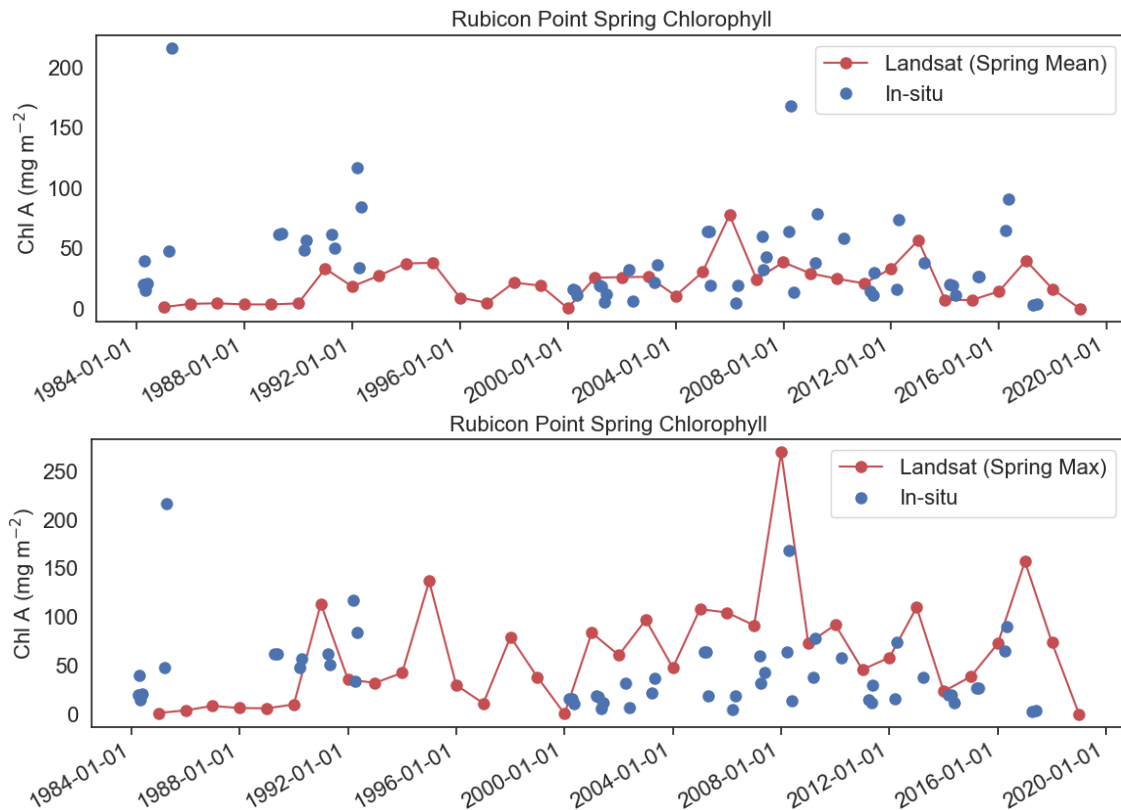


Figure 7: Landsat estimated spring (March-May) periphyton Chl A at Rubicon Point based on multivariate regression using the SWIR1, Red, and Green bands. Mean (Top) and maximum (Bottom)

springtime (March-May) periphyton chlorophyll-a at the Rubicon Point site based on Landsat estimated values.

In addition to exploring the potential to estimate the absolute concentration of periphyton, we explored the potential to estimate relative concentrations with respect to historical highs and lows for specific periods of record. Recent advances in cloud computing applications such as Climate Engine (Huntington et al., 2017) has created the ability for researchers and managers alike to produce anomaly maps, where the spatial distribution of water quality metrics can be assessed relative to the long-term average spatial distribution for a specific date or range of dates. Assessing relative concentrations through the calculation of spatial or temporal anomalies is useful given that there is uncertainty in measuring or predicting absolute concentrations, and there is also uncertainty in in-situ measurements with respect to how spatially and temporally representative those measurements are (due to variable sampling locations, and high degree of temporal variability as illustrated in Figure 6). To illustrate this point, an anomaly map of the NASA OC2 predicted mean chlorophyll-a was computed for May-June of 2014 using Climate Engine (Figure 8). A widely observed and publicized algae bloom during this time occurred near Regan Beach located in Southlake Tahoe (Figure 8; Left). The predicted high OC2 anomaly corresponded with public observations of the algae bloom (Figure 8; Right). While absolute concentrations will be uncertain, it is argued here that anomaly mapping has potential utility for monitoring relative concentrations of chlorophyll-a in both time and space.

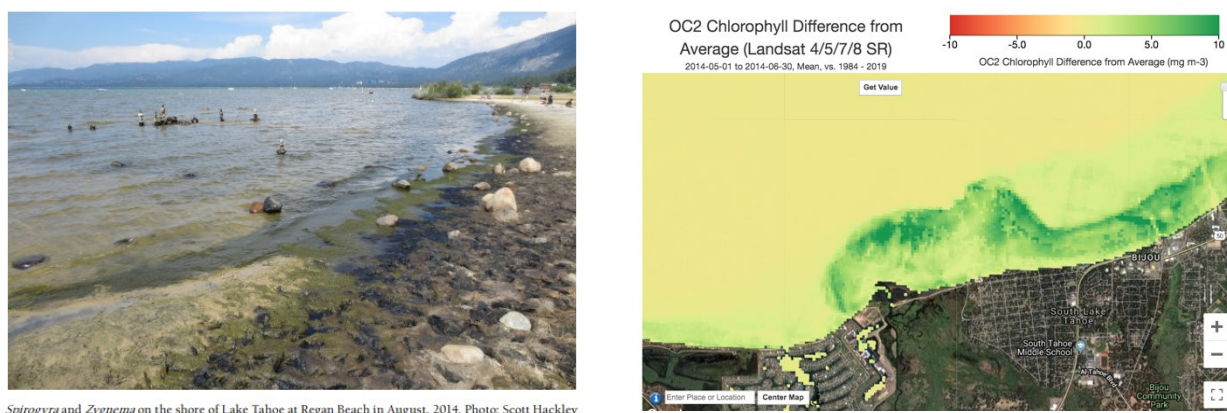


Figure 8: Observation of a well-known algae bloom during spring of 2014 near Regan Beach, South Lake Tahoe (left), and the corresponding anomaly map of the OC2 predicted chlorophyll-a using ClimateEngine.org. Green coloring along the lake's shoreline corresponds with higher than average algae concentrations.

Spring synoptic sampling at Kiva Point near South Lake, CA reported both high and low concentrations of chlorophyll-a over a 13-year sampling period. Anomaly maps of the green reflectance computed with Climate Engine for select years illustrate that periods of anomalously high and low green reflectance correspond well to periods of anomalously high and low in-situ periphyton chlorophyll-a. These results again highlight the usefulness of satellite-based

anomalies for monitoring and better understanding current conditions relative to long-term average conditions for specific locations and time periods.

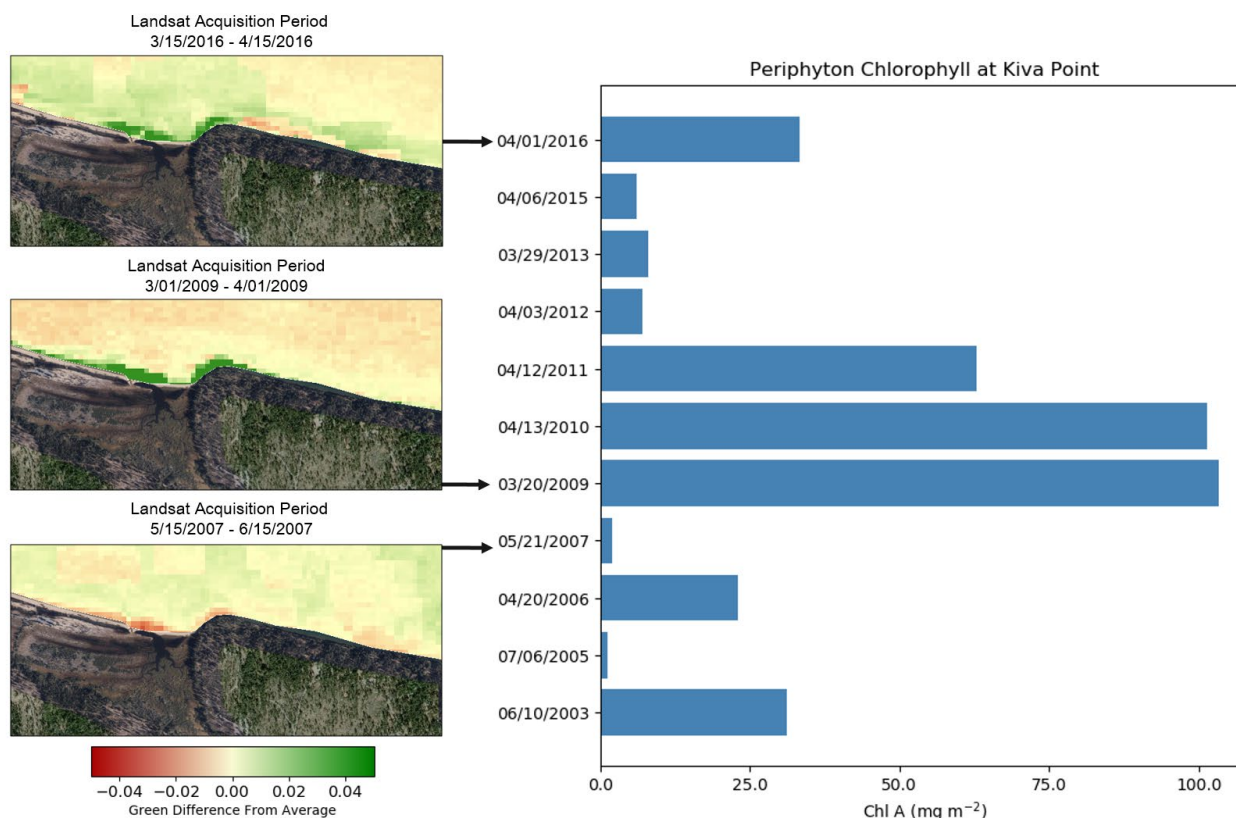


Figure 9: Spring synoptic periphyton Chl A time series at Kiva Point near South Lake, CA beside Green Band anomaly maps for periods of low, mid, and high concentrations. Green pixels represent higher than average green reflectance, while red pixels represent lower than average green reflectance. Climate Engine Links for above maps: Bottom, <https://climengine.page.link/eK8e>; Center: <https://climengine.page.link/Haik>; Top: <https://climengine.page.link/Q48M>

Summary and Recommendations

The purpose of this study was to broadly assess the potential for the use of satellite imagery that is atmospherically corrected using standard algorithms and is freely available for operational monitoring of nearshore periphyton at Lake Tahoe. Results from this study indicate that the use of Landsat satellite imagery for nearshore periphyton chlorophyll-a monitoring shows promise, however, site specific algorithms should be developed for producing acceptable results. Although Landsat-based remote sensing may not be able to provide skillful predictions of chlorophyll-a for all sites throughout the lake, it shows utility for monitoring periphyton growth for specific sites and seasons. Landsat's unparalleled historical coverage (1984-present) make it invaluable for understanding long-term variability of nearshore water quality at Lake Tahoe. Test

applications of multivariate regression at Incline West and Rubicon Point shows utility for predicting inter-annual variability of chlorophyll-a observed in the routine sample record, and is potentially useful for gap filling and completing the historical record at sites with sufficient data coverage. The in-situ periphyton chlorophyll-a dataset shows large variability at short time scales (2 weeks to monthly), making short-term trends difficult to interpret. Anomaly mapping of periphyton chlorophyll-a was shown to correspond well with in-situ anomalies, and provides useful information about relative concentrations in space and time. New developments and applications in remote sensing and machine learning show promise for more robust predictions where more conventional statistical approaches applied in this work fail. While our cursory statistical assessment of Landsat satellite and in-situ periphyton concentrations conducted in this study show some predictive skill using traditional statistical methods, the use of in-situ periphyton dataset for training machine learning models should be explored as follow on work to this study.

Recommendations

We recommend a collaborative review of existing data collection methods in the context of historic and future satellite image acquisitions. Specific areas for improvement could include:

- Documenting precise GPS locations for all in-situ sampling
- Targeting satellite overpass dates for in-situ sampling to improve the temporal alignment between satellite and in-situ sampling
- Developing GIS polygons that represent potential sample area zones for improved spatial representation and averaging of satellite data
- In-situ spectral monitoring to complement and refine satellite image surface reflectance estimates.
- Further statistical modeling by season and site using multiple linear regression and the inclusion of new satellite data (i.e. Sentinel 2 and 3)
- Use of in-situ periphyton data for training of satellite image based machine learning models (such as TensorFlow), and assessment of predictive skill of these higher level statistical approaches

We recommend additional analyses similar to those performed in this study using new European Space Agency (ESA) operational satellite data such as Sentinel 2a, 2b, and 3. Sentinel 2a,b are satellite missions that are “Landsat like” but have higher spatial and temporal resolution. Combining Sentinel 2a and 2b yields a 5-day revisit, at 10 m spatial resolution for red, green, and blue bands. Combining Sentinel 2 with Landsat 8 at 30m resolution yields a revisit of ~3 days. Sentinel 2a was launched in June of 2015 and Sentinel 2b was launched in March of 2017. Sentinel 3a was launched in February of 2016, and Sentinel 3b was launched in April of 2018. Sentinel 3 is primarily an ocean mission, however, there are numerous potential land and inland water applications. The revisit of Sentinel 3 is approximately 2 days, and the spatial resolution is 300 m. Because Sentinel 3 was primarily designed as an ocean application mission, its spectral resolution is high compared to Landsat and Sentinel 2. Sentinel 3’s Ocean and Land Colour Instrument (OLCI) measures reflectance in 21 bands (Landsat 8 measures reflectance in 11 bands) and is currently the only sensor in space able to detect cyanobacteria. The US EPA

recently developed an operational platform for monitoring cyanobacteria and other harmful algal blooms based on Sentinel 3 data (Schaeffer et al., 2018). Sentinel 2 and 3 data are available in Google Earth Engine and could easily be integrated into a future study that builds off and leverages software developed in this study. The potential to integrate Landsat with Sentinel 2 and 3 data and improve water quality monitoring potentials at Lake Tahoe is promising and should be investigated further. In addition, the potential for application of machine learning models should be explored using Landsat and Sentinel data along with in-situ periphyton data to train these models. Models that can easily be applied to satellite data such as TensorFlow could be explored relatively quickly given that Google Earth Engine has integrated TensorFlow Neural Network functions (<https://developers.google.com/earth-engine/tensorflow>) and our team has specialized expertise and close connections with Google developers.

In summary, the use of remote sensing is a feasible option for monitoring nearshore water quality at Lake Tahoe, however, there are challenges. Integration and use of new satellite data combined with targeted sampling aligned with satellite image acquisition dates could yield long-term benefits for monitoring with satellite remote sensing. While other water quality remote sensing platforms are available, such as the use of piloted aircraft and drones, these platforms are expensive and extremely limited for long term operational monitoring and applications. We believe that investing in research and applications that combine in-situ data with freely available satellite remote sensing data is the best option for ensuring affordable, sustainable, and scalable long-term monitoring capabilities at Lake Tahoe.

Acknowledgments

This work was supported by funding from the Nevada Division of Environmental Protection. We thank The University of California-Davis and Tahoe Environmental Research Center for collection and sharing of the historic nearshore periphyton dataset.

References

- Allan, M. G., Hamilton, D. P., Hicks, B. J., & Brabyn, L. (2011). Landsat remote sensing of chlorophyll a concentrations in central North Island lakes of New Zealand. *International Journal of Remote Sensing*, 32(7), 2037-2055.
- Balali, S., Hoseini, S. A., Ghorbani, R., & Balali, S. (2012). Correlation of chlorophyll-A with Secchi Disk Depth and Water Turbidity in the International Alma Gol Wetland, Iran'. *World J. Fish Marine Sci*, 4, 504-508.
- Barrett, D. C., & Frazier, A. E. (2016). Automated Method for Monitoring Water Quality Using Landsat Imagery. *Water*, 8(6), 257.
- 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(4), 373-382.
- Concha, Javier A. (2015). The Use of Landsat 8 for Monitoring of Fresh and Coastal Waters. Thesis. Rochester Institute of Technology. Accessed from <http://scholarworks.rit.edu/theses/8895>

- Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay. Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, **12**, 2825-2830 (2011) ([publisher link](#))
- Glooschenko, W. A., & Blanton, J. O. (1977). Short-term variability of chlorophyll a concentrations in Lake Ontario. *Hydrobiologia*, 53(3), 203-212.
- 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*.
- Heyvaert, A. C., Reuter, J. E., Chandra, S., Susfalk, R. B., Schladow, S. G., Hackley, S. G. (2013). Lake Tahoe Nearshore Evaluation and Monitoring Framework. Final report prepared for the USDA Forest Service Pacific Southwest Research Station.
- Ichimura, S. (1960). Diurnal fluctuation of chlorophyll content in lake water. *Bot. Mag. Tokyo*, 73, 217-224.
- John D. Hunter. Matplotlib: A 2D Graphics Environment, *Computing in Science & Engineering*, **9**, 90-95 (2007), DOI:10.1109/MCSE.2007.55 ([publisher link](#))
- K. Jarrod Millman and Michael Aivazis. Python for Scientists and Engineers, *Computing in Science & Engineering*, **13**, 9-12 (2011), DOI:10.1109/MCSE.2011.36 ([publisher link](#))
- Kloiber, S. M., 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(1), 38-47.
- Mohamed, M. F. (2015). Satellite data and real time stations to improve water quality of Lake Manzalah. *Water Science*, 29(1), 68-76.
- NASA Goddard Space Flight Center, Ocean Biology Processing Group. (2014). Sea-viewing Wide Field-of-view Sensor (SeaWiFS) Ocean Color Data, NASA OB.DAAC, Greenbelt, MD, USA. http://doi.org/10.5067/ORBVIEW-2/SEAWIFS_OC.2014.0. Accessed 2016/10/31. Maintained by NASA Ocean Biology Distributed Active Archive Center (OB.DAAC), Goddard Space Flight Center, Greenbelt MD.
- Pahl, R., Huntington, J. L. (2010). Assessment of Remote Sensing Techniques for the Characterization of Lake and Reservoir Clarity, Nevada Water Resources Association 2010 Annual Conference: Las Vegas
- Rosen, M.R., K. Turner, S. Goodbred and J.M. Miller. 2012. A synthesis of aquatic science for management of Lakes Mead and Mohave. U.S. Geological Survey Circular 1381. 162 pp.
- Schaeffer, B. A., Bailey, S. W., Conmy, R. N., Galvin, M., Ignatius, A. R., Johnston, J. M., ... & Urquhart, E. A. (2018). Mobile device application for monitoring cyanobacteria harmful algal blooms using Sentinel-3 satellite Ocean and Land Colour Instruments. *Environmental modelling & software*, 109, 93-103.
- Schmidt, G., Jenkerson, C., Masek, J., Vermote, E., & Gao, F. (2013). Landsat ecosystem disturbance adaptive processing system (LEDAPS) algorithm description (No. 2013-1057). US Geological Survey.

- Steissberg, T., Schladow, G., Hook, S. Monitoring Past, Present, and Future Water Quality Using Remote Sensing (2010). Report prepared for the Southern Nevada Public Lands Management Act; Lake Tahoe Environmental Improvement Program.
- Stéfan van der Walt, S. Chris Colbert and Gaél Varoquaux. The NumPy Array: A Structure for Efficient Numerical Computation, *Computing in Science & Engineering*, **13**, 22-30 (2011), DOI:10.1109/MCSE.2011.37 ([publisher link](#))
- Tahoe Environmental Research Center. 2017. Lake Tahoe Water Quality Investigations: Algal Growth Potential Assays, Phytoplankton, Periphyton. Annual report 6/30/2017. <https://tahoe.ucdavis.edu/sites/g/files/dgvnsk4286/files/inline-files/2017%20TERC%20Annual%20Rpt.pdf> Accessed: 9/13/2019.
- Travis E. Oliphant. Python for Scientific Computing, *Computing in Science & Engineering*, **9**, 10-20 (2007), DOI:10.1109/MCSE.2007.58 ([publisher link](#))
- Travis E. Oliphant. A guide to NumPy, USA: Trelgol Publishing, (2006).
- Torbick, N., Hu, F., Zhang, J., Qi, J., Zhang, H., & Becker, B. (2008). Mapping chlorophyll-a concentrations in West Lake, China using Landsat 7 ETM+. *Journal of Great Lakes Research*, 34(3), 559-565.
- USGS (2018a). Landsat Surface Reflectance Level-2 Science Products. <https://landsat.usgs.gov/landsat-surface-reflectance-data-products>
- U.S. Geological Survey. (2018b). Landsat 8 Surface Reflectance Code (LaSRC) Product, version 4.3, March 2018, 40 p.
- Waxter, M. T. (2014). Analysis of Landsat Satellite Data to Monitor Water Quality Parameters in Tenmile Lake, Oregon.
- Wes McKinney. Data Structures for Statistical Computing in Python, *Proceedings of the 9th Python in Science Conference*, 51-56 (2010) ([publisher link](#))
- Wrobel, C. A., and Carlos Alberto Brebbia, eds. (2012). Water pollution: Modelling, measuring and prediction. Springer Science & Business Media.
- Zhu, Z., and Woodcock, C. E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote sensing of environment*, 118, 83-94.