



Hindcasting water quality in an optically complex system

C. Hyatt Hansen¹, P. Dennison¹, S. Burian¹, M. Barber¹
& G. Williams²

¹*University of Utah, USA*

²*Brigham Young University, USA*

Abstract

As is the case with many large lakes, field sampling records (and the understanding of historical water quality) in the Great Salt Lake natural surface water system are heavily limited due to time and cost constraints, as well as a number of independent organizations collecting and managing data. To address these deficiencies, remote sensing of surface water quality is used to hind-cast historical conditions of algal blooms in the GSL surface water system (GSLSWS). This system is unique because its lakes are closely connected, yet have widely varying characteristics and conditions. An approach for development of lake-specific models is demonstrated, using historical Landsat and field-sampled data. This study builds on previous studies of historical water quality which have used broad-spectral remote sensing data and near-coincident field samples by evaluating the ability of models to accurately estimate water quality under optically complex conditions (such as high turbidity and shallow conditions). We also examine the spatiotemporal variability of the field-sampled data and address the issue of near-coincidence between a historic dataset of field-samples and remote sensing images. Existing field sampling campaigns for this area do not provide sufficient information about adverse conditions or long-term spatiotemporal patterns. Results of the remote sensing model application however may provide useful metrics for algal bloom dynamics, including timing of blooms, duration and spatial extent of algal blooms, and how these dynamics vary over time and within the surface water system. Products of the remote sensing models broaden the foundation of understanding of water quality conditions, which can be used to move forward with better monitoring and management practices.

Keywords: remote sensing, Great Salt Lake, chlorophyll, turbidity, Landsat.



1 Introduction

The Great Salt Lake surface water system (GSLWS) consists of the Great Salt Lake, Utah Lake, and the Jordan River, which flows from Utah Lake to the Great Salt Lake. The system offers an excellent opportunity to examine a large surface water system that is closely connected with a rapidly developing urban area. The GSLWS is highly unique in the degree of urban influence and management of the system, the physical and chemical characteristics of the inland lakes, and the interconnected relationship of the system's distinct water bodies. The GSLWS begins with a series of streams and reservoirs in the Wasatch Mountains which flow west through the urban area of Salt Lake City-Provo, Utah and discharge into the Great Salt Lake and Utah Lake. Water from Utah Lake flows north through the Jordan River, which also receives discharge from the urban area before it terminates in the southern end of Farmington Bay. Each of the bays of the Great Salt Lake (Gilbert, Farmington, Gunnison, and Willard/Bear River) are considered as distinct lakes, due to both natural divisions and man-made causeways which result in distinct characteristics in each of the bays. For instance, the southern bays are much less saline than the northern bays and are generally deeper than the northern bays due to restricted flow through breaches in the east-west causeway.

The most heavily trafficked portions of the lake (for recreation and for wildlife habitat) are these southern bays and Utah Lake. Farmington Bay supports an ecosystem that is of hemispheric importance for millions of migratory birds who come to nest and feed and on the abundant insects and brine shrimp. It is also a popular recreation spot for hunters, bird-watchers, and outdoor enthusiasts. Gilbert Bay plays an important role in both the brine shrimp and recreation industries. Utah Lake is an important year-round habitat for waterfowl and a number of fish, including the endangered June Sucker. Like the southern half of the GSL, Utah Lake is also an important staging location for many migratory birds. Thus, for the purpose and focus of this study, the GSLWS is limited to this portion of the system (as highlighted in Figure 1).

1.1 Harmful algal blooms in the GSLWS

A major water quality concern in the GSL system is the amount of algae in the lakes. While algae are a key component of the food web, excessive amounts of algae or harmful algal blooms (HABs) can negatively affect the lake ecosystem. HABs can be either photosynthetic algae whose cellular structures or collective biomass negatively affect food web dynamics or protozoans with the ability to produce toxins [1]. It is important to note that not all species (such as some diatoms with hard outer shells) are able to be digested by the shrimp and insects. These non-beneficial algae mainly contribute to habitat/aesthetic problems caused by algal decay and depletion of dissolved oxygen as bacteria decompose the algae [2, 3].



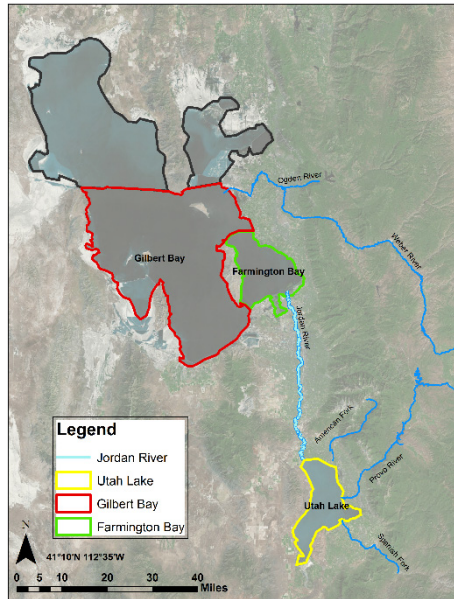


Figure 1: Overview of Great Salt Lake Surface Water System. For the purpose of this study, the system is limited to the lakes outlined in colour.

Additionally, one of the most common species of algae in Farmington Bay is a toxic cyanobacteria, *Nodularia* [3, 4]. Cyanobacteria have been shown to cause illnesses and rashes on swimmers in several Utah lakes, and they have even been linked to several canine deaths [3, 5]. Both toxic and non-toxic HABs have impacted the GSLSWS by negatively affecting lake aesthetics [6]. With bird-watching, fishing, and recreation drawing thousands of visitors to the lakes each year, there is considerable motivation to study the conditions that are impacting lake aesthetics, ecosystem health, and posing risks to human/animal health.

1.2 Current and alternative monitoring of HABs in the lake system

Despite the motivation for studying algal blooms in the GSLSWS, there is a relatively limited understanding of algal bloom dynamics (such as timing, spatial extent, or overall magnitude) and the factors that contribute to these dynamics. A number of studies, based on concentrated field sampling campaigns, have suggested connections between algal population dynamics and in-lake characteristics such as salinity, nutrients, water temperature, etc., [4, 7–11]. Water resource managers in the region are especially interested in understanding how external factors, such as changing climate and hydrologic conditions, and urbanization of local watersheds [12, 13] contribute to poor lake conditions and HABs. However, this is significantly inhibited due to the limited and sporadic nature of the historical field record, as well as differences in data management and sharing techniques between organizations that monitor water quality in the

region. Improving the current understanding will require alternative forms of monitoring and representing the system. In recent decades, remote sensing models have been used to overcome the limitations associated with traditional water quality monitoring and assessment techniques. Remote sensing methods rely on the fact that natural surfaces reflect energy (radiance) from the sun back into space, which can be measured remotely by a satellite-mounted sensor or by an airborne camera. Reflected radiance is wavelength-dependent, and the amount of reflectance at each wavelength can be influenced by the absorption of pigments in algae and other constituents. This results in distinct spectral signatures for different lake conditions, such as pure water, water with algae, or water with suspended sediments. The spectral signature is influenced by the amount of chlorophyll-a (chl-a) – as well as other pigments and in-lake constituents - which allows for detection of approximate concentrations.

Over the past 35 years, remote sensing applications for water quality have been demonstrated using a range of instruments, locations, modelling complexity, and application scope. However, relatively few studies [14, 15] exploit historical records to develop models or examine long-term trends, and many water quality monitoring agencies do not utilize the potential information. This study demonstrates simple approach for developing and applying remote sensing using an existing historical field sampling record and Landsat (due to its long-term continuous record and availability of processed data products). Additionally, we evaluate this approach in light of the limitations of historical datasets and the unique study area of the GSLSWS. Enhancing the long-term historic water quality record of the GSLSWS through remote sensing will help water quality managers understand the system more holistically.

1.3 Challenges associated with remote sensing of Chl-a

Like many lake systems, the irregular nature of the historic record for the study area results in few exact matches between field sampling and image acquisition dates. In ocean/lake water quality remote sensing literature, there is a range of time-windows for considering data to be a “match.” Depending on the study, these windows range from a difference of ± 3 hours [16], 1 day [17], 7 days [18, 19], to ± 10 days [14] between the satellite image acquisition and the field sample used for calibration. Often, a particular time-window for near-coincident matches is arbitrarily chosen (e.g. using an arbitrary increase in the percentage of samples that match with a satellite image [20]) rather than demonstrating an improved model performance to justify the use additional near-coincident data or actual short-term temporal variability in the field record.

Additionally, there may be potential challenges to remote sensing models in the GSLSWS caused by the reflectance of the bottom surface and high turbidity (particularly in Utah Lake) which can interfere with the spectral signature of the surface water. Several studies have used hyperspectral sensors to account for these interference effects [21] and others have used non-linear regression techniques or complex spectral un-mixing techniques with broad-band sensors [22, 23].



2 Data and methods

2.1 Description of data

Landsat surface reflectance imagery data are provided by the USGS and hosted on the Google Earth Engine data catalogue. These data are generated from Landsat-5, 7, and 8 imagery which has been processed using Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) software to provide surface reflectance. The use of surface reflectance data helps eliminate the effects atmospheric interferences such as aerosols that may be present in top-of-atmosphere reflectance products.

In-situ samples of chl-a (corrected for pheophytin) were collected by the United States Geological Survey (USGS) and the Jordan River/Farmington Bay Water Quality Council for the Great Salt Lake, and by the Utah Division of Water Quality (UDWQ) for Utah Lake. The number of available data, given the relaxation of the time-windows, is summarized in Table 1.

Table 1: Number of near-coincident “matches” by time-window.

Time-Window	Farmington Bay	Gilbert Bay	Utah Lake
Same Day	5	54	6
±1 day	12	165	24
±2 days	22	372	57
±3 days	26	429	65
±4 days	34	456	72
±5 days	34	491	72
±6 days	45	544	92
±7 days	45	580	95

2.2 Variability in the historic record

As mentioned previously, the field record for the study area is highly limited in the information it can provide about spatial and/or temporal variability because of irregular and infrequent revisit times. shows the variability in chl-a concentrations as a percent difference between subsequent samples. Very few samples (at the same location) in the historical record were collected within a few days of each other, so the true variability over short time periods is not easily discerned from this dataset.

Thus, despite a large dataset of historical samples, the historical datasets alone cannot be used to determine an appropriate time-window for near-coincident matches. Instead, a number of time-windows from ±1 to 7 days of the *in situ* sample collection date are considered for model development.



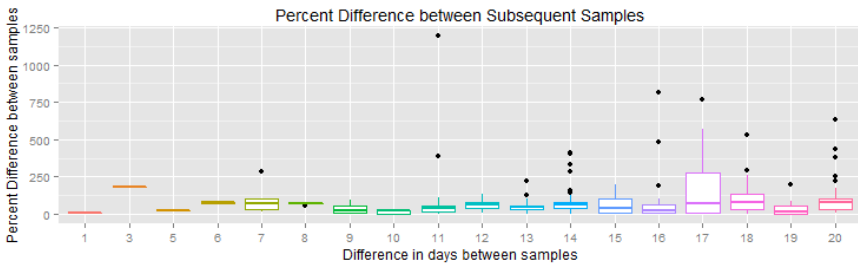


Figure 2: Temporal variability in Chl-a magnitudes of subsequent samples.

2.3 Model specification

The process of imagery data-retrieval and formatting was automated through a python script and the *ee* (Google Earth Engine) Python package. The program connects to the Google Earth Engine data catalogue, filtered imagery data from within the specified time-window, and retrieved reflectance data from pixels within 1 meter of the corresponding field sample location. The python script and specifically the use of the *ee* package facilitated rapid data retrieval and eliminated the need for storing and managing large imagery datasets on local computers or servers [24]. The USGS surface reflectance product also contains additional bands for a cloud mask and level of confidence for the cloud mask, which eliminates the need to manually verify whether the pixel is free of clouds and shadows.

Empirical models were then developed for time-windows ranging from 0 to ± 7 days. First, stepwise regressions were used to determine significant parameters from all bands, band ratios, and indices from literature [25, 26], and for Gilbert Bay, a seasonal indicator variable for June-September when chl-a values are lower in magnitude and less variable. A generalized linear modelling approach with a log-link function was used to fit the data and account for non-normality of the dependent variable.

Each of the lake-specific models were evaluated for model stability and predictive ability through a k-fold ($k=5$) cross validation technique and comparing the model R^2 and the average R^2 from the k-folds.

Model predictive ability in optically complex conditions (shallow depth and high turbidity) was evaluated by comparing the RMSE for a subset of the samples in which the depth of the water column was recorded and where turbidity was also measured at the same location and date as chl-a.

3 Results

3.1 Evaluating near-coincident data

While other studies report an increase in model descriptive ability (reported as R^2) as the time-window is increased, the GLSWS data do not demonstrate the

same behaviour. shows that, generally, as the time-window is increased, the descriptive ability of the model does not increase. For Farmington Bay and Utah Lake, where there are considerably fewer potential matches, there is a local maximum in model performance at ± 4 days, which may indicate this to be an appropriate choice for model development and application.

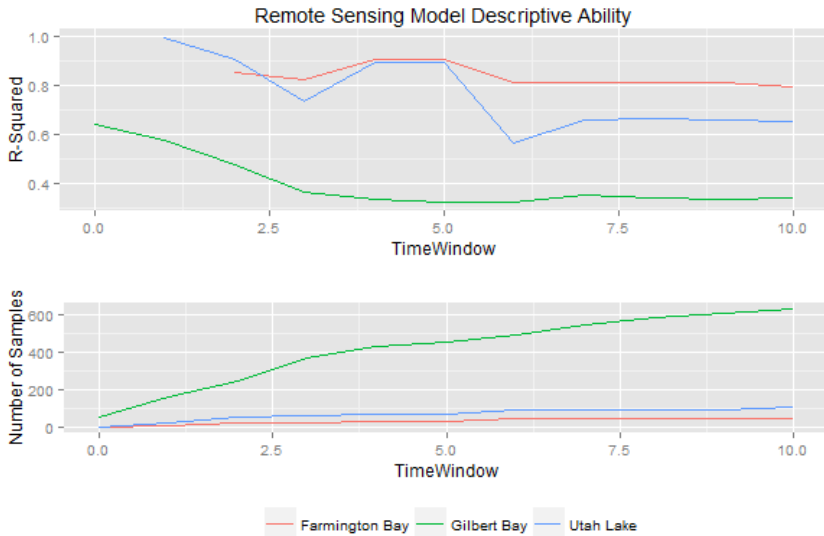


Figure 3: Coefficient of determination (R^2) and number of near-coincident samples per time window.

3.2 Performance in highly turbid and shallow conditions

To evaluate the performance of the models in optically complex conditions, subsets of the data were created that corresponded with additional measurements of lake depth (available for Gilbert Bay), and water clarity (for Utah Lake), as measured by total suspended solids (TSS) and Secchi depth. Data was further subset using percentiles to represent poor optical conditions. Independent t-tests indicate that the mean of the model errors for samples in the lowest 25% of the lake depths in Gilbert Bay is not statistically significantly different from the mean of the errors for samples in deeper parts of the lake (p -value=0.18). For Secchi depth, the mean in model errors was not significantly different between samples corresponding with the bottom 25% and samples with greater Secchi depth measurements (p -value=0.15). For TSS however, the t-test indicated a significant (p -value=0.01) difference in mean model errors for highly turbid (>75 percentile) areas of Utah Lake.

3.3 Potential model application

One of the primary benefits for using historical imagery data and remote sensing is the ability to estimate historical conditions where there are no other data available. To demonstrate, the model for Farmington Bay was applied to a median-filtered image for the month of September from 1990–1994 using the Google Earth Engine API. Estimates for two locations, near Antelope Island, and the other near the inlet south of the causeway are shown in .

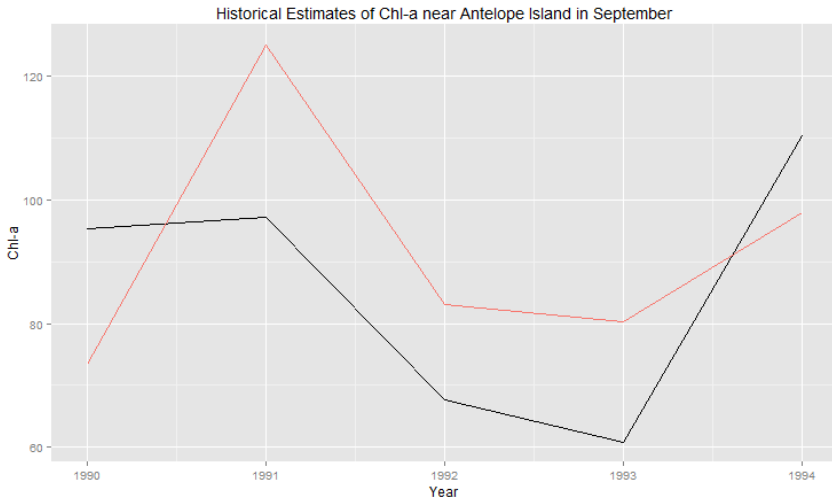


Figure 4: Estimates of Chl-a for two locations in Farmington Bay. Red: near the inlet; black: near Antelope Island.

While there are no samples in the historical record from 1990–1994, estimates of chl-a at individual locations (or regions) can be obtained. The simple example illustrated in Figure 4 demonstrates a case where monitoring agencies may be interested in how conditions near freshwater inflows to the lake compare to other areas, including areas near popular recreational spots. In this case, it appears that the area near the inlet generally experiences higher levels of chl-a than areas near the island.

4 Discussion and conclusion

Despite the limitations in the existing field record, near-coincident data has been shown to be useful in development of chl-a estimation algorithms. According to the datasets and model development techniques demonstrated in this study, shorter time-windows were found to be generally more appropriate as they produced models with higher descriptive ability. While interferences from shallow waters and high turbidity may pose an issue to remote sensing applications in this area, preliminary analysis indicates that the remote sensing models may perform as well in these complex waters as in other areas.

The remote sensing models developed in this study demonstrate the utility of exploiting available datasets (both imagery and historic field samples) to gain additional information and enhance understanding of long-term trends in a complex body of water. The temporal trends provided by the remote sensing models provide a more complete description of historical conditions in the lake system. This allows for comparison in different parts of the lake that have not been monitored in the past, as well as for more frequent estimates for areas that are monitored to fill in gaps in the historical record. Additionally, water quality and natural resource professionals may be interested in other algal bloom dynamics that affect recreation and wildlife habitat include the timing of high magnitudes, the overall conditions of the lake, and the spatial extent of the blooms. These measurements, while not available through the historical field sampling record, can be obtained through remote sensing estimates and supplement current monitoring and evaluating efforts.

References

- [1] Anderson, D.M., P.M. Glibert, and J.M. Burkholder, *Harmful algal blooms and eutrophication: nutrient sources, composition, and consequences*. Estuaries, 2002. **25(4)**: pp. 704-726.
- [2] Marcarelli, A.M., Mills, Michael D., Wurtsbaugh, Wayne A., *The Great Salt Lake Doesn't Stink...But Farmington Bay Does!* 2001.
- [3] Wurtsbaugh, W.A., *Nutrient Loading and Eutrophication in the Great Salt Lake*. Watershed Sciences Faculty Publications, 2008.
- [4] Wurtsbaugh, W.A., A.M. Marcarelli, and G.L. Boyer, *Eutrophication and Metal Concentrations in Three Bays of the Great Salt Lake (USA)*. 2009 Final Report to the Utah Division of Water Quality, Salt Lake City, Utah, 2012: p. 1.
- [5] Penrod, E., *After Utah Lake-related dog deaths, experts recommend protocol for toxic algae*, in *Salt Lake Tribune*. 2015: Salt Lake City, UT.
- [6] Nicholson, B., and Marcarelli, A., *The Paradox of a Great Salt Lake*, in *Southwest Hydrology*. 2004, Department of Aquatic, Watershed and Earth Resources, Utah State University. pp. 24-25.
- [7] Wurtsbaugh, W.A. and T.S. Berry, *Cascading effects of decreased salinity on the plankton chemistry, and physics of the Great Salt Lake (Utah)*. Canadian Journal of Fisheries and Aquatic Sciences, 1990. **47(1)**: pp. 100-109.
- [8] Goel, R. and L. Myers, *Evaluation of Cyanotoxins in the Farmington Bay, Great Salt Lake, Utah*. Project Report. Central, 2009.
- [9] McCulley, E.W., Wayne; Barnes, Brian, *Factors affecting the spatial and temporal variability of cyanobacteria, metals, and biota in the Great Salt Lake, Utah*. 2015.
- [10] Marden, B.M., Miller, T., Richards, D., *Factors Influencing Cyanobacteria Blooms in Farmington Bay, Great Salt Lake, Utah*, in *A Progress Report of Scientific Findings From the 2013 Growing Season*. 2015, The Jordan River/Farmington Bay Water Quality Council.
- [11] Merritt, L.B., *Utah Lake: A Few Considerations*. 2014.



- [12] Whitehead, P., *et al.*, *A review of the potential impacts of climate change on surface water quality*. Hydrological Sciences Journal, 2009. **54(1)**: pp. 101-123.
- [13] Naftz, D., *et al.*, *Anthropogenic influences on the input and biogeochemical cycling of nutrients and mercury in Great Salt Lake, Utah, USA*. Applied Geochemistry, 2008. **23(6)**: pp. 1731-1744.
- [14] Olmanson, L.G., M.E. Bauer, and P.L. Brezonik, *A 20-year Landsat water clarity census of Minnesota's 10,000 lakes*. Remote Sensing of Environment, 2008. **112(11)**: pp. 4086-4097.
- [15] Hansen, C.H., G.P. Williams, and Z. Adjei, *Long-Term Application of Remote Sensing Chlorophyll Detection Models: Jordanelle Reservoir Case Study*. Natural Resources, 2015. **6(02)**: p. 123.
- [16] Bailey, S.W. and P.J. Werdell, *A multi-sensor approach for the on-orbit validation of ocean color satellite data products*. Remote Sensing of Environment, 2006. **102(1)**: pp. 12-23.
- [17] Lesht, B.M., R.P. Barbiero, and G.J. Warren, *A band-ratio algorithm for retrieving open-lake chlorophyll values from satellite observations of the Great Lakes*. Journal of Great Lakes Research, 2013. **39(1)**: pp. 138-152.
- [18] McCullough, I.M., C.S. Loftin, and S.A. Sader, *Landsat imagery reveals declining clarity of Maine's lakes during 1995–2010*. Freshwater Science, 2013. **32(3)**: pp. 741-752.
- [19] Kloiber, S.M., *et al.*, *A procedure for regional lake water clarity assessment using Landsat multispectral data*. Remote Sensing of Environment, 2002. **82(1)**: pp. 38-47.
- [20] Johnson, R., *et al.*, *Three improved satellite chlorophyll algorithms for the Southern Ocean*. Journal of Geophysical Research: Oceans, 2013. **118(7)**: pp. 3694-3703.
- [21] Cannizzaro, J.P. and K.L. Carder, *Estimating chlorophyll a concentrations from remote-sensing reflectance in optically shallow waters*. Remote Sensing of Environment, 2006. **101(1)**: pp. 13-24.
- [22] Liu, Y., M.A. Islam, and J. Gao, *Quantification of shallow water quality parameters by means of remote sensing*. Progress in Physical Geography, 2003. **27(1)**: pp. 24-43.
- [23] Tyler, A., *et al.*, *Remote sensing of the water quality of shallow lakes: A mixture modelling approach to quantifying phytoplankton in water characterized by high-suspended sediment*. International Journal of Remote Sensing, 2006. **27(8)**: pp. 1521-1537.
- [24] Hansen, C.H., *Google Earth Engine as a Platform for Making Remote Sensing of Water Resources a Reality for Monitoring Inland Waters*.
- [25] Fuller, L.M., S.S. Aichele, and R.J. Minnerick, *Predicting water quality by relating Secchi-disk transparency and chlorophyll a measurements to satellite imagery for Michigan inland lakes, August 2002*. 2004: US Department of the Interior, US Geological Survey.
- [26] Panda, S., V. Garg, and I. Chaubey, *Artificial neural networks application in lake water quality estimation using satellite imagery*. Journal of Environmental Informatics, 2004. **4(2)**: pp. 65-74.