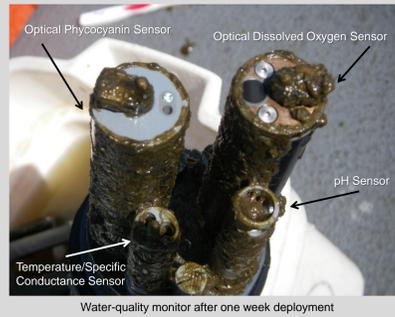


In Cooperation with the Bureau of Reclamation

# Water Quality Dynamics And Phycocyanin Detection As A Biomass Indicator In Upper Klamath Lake, Oregon, 2011

D.B. Eldridge<sup>1</sup>, L.Schenk<sup>1</sup>, S.L. Caldwell Eldridge<sup>1</sup>, T.M. Wood<sup>2</sup>, and S. Wherry<sup>2</sup>

U.S. Geological Survey: <sup>1</sup>Klamath Falls Field Station, Klamath Falls, OR 97603, <sup>2</sup>Oregon Water Science Center, Portland, OR 97201



Cyanobacteria bloom, Upper Klamath Lake, OR

## Introduction and Purpose

Poor water quality in Upper Klamath Lake, a large (surface area 232 km<sup>2</sup>), shallow (average depth 2.8 m), hypereutrophic water body in south-central Oregon, results from the growth and decomposition of seasonal blooms of dominantly *Aphanizomenon flos-aquae* (AFA) and contributes to the decline of endangered, Lost River (*Delistes luxatus*) and shortnose (*Chasmistes brevirostris*) suckers (Perkins et al., 2000; Banish et al., 2009).

Recent studies have demonstrated the use of empirically derived regression equations to estimate water-quality constituents based on continuously measured physical and chemical data (Christensen, Jian, and Ziegler, 2000; Ryberg, 2006; Wood and Gartner, 2010). This "surrogate" approach has been used in rivers to estimate the concentration and, with stream gage data, the load (concentration multiplied by discharge) of water-quality constituents that cannot be measured continuously. In lakes, transport is not uni-directional, so the load of a constituent is not meaningful, but the quantity of interest is often the storage, or whole-lake integrated mass of a constituent. In Upper Klamath Lake, the storage of biomass and nutrients is used for calculating lake mass balances (Kann and Walker, 2001) and for calibrating and validating models to establish total maximum daily loads (Walker, 2001). Storage has previously been calculated from water-quality profiles collected at twice-monthly or monthly intervals (Kann and Walker, 2001).

The purpose of this study is to use continuously measured data from multi-parameter water-quality monitors (sondes) to calculate daily depth-averaged measures of biomass at two sites as a first step in demonstrating the feasibility of calculating whole-lake storage using the same technique at additional sites. The dependent variables were total phosphorus and chlorophyll *a* concentrations measured from depth-averaged water samples collected weekly and cyanobacteria cells, quantified as relative fluorescent units (IRFU) per square meter and determined from a numerical integration of weekly phycocyanin depth profiles.

## Methods

### Data Collection

Water samples were collected weekly from May to October 2011 at sites MDN and MDT (fig. 1) and analyzed for total nitrogen, total phosphorus, and chlorophyll *a*. Samples were also collected at three-hour intervals over 24 hours twice during moderate-to-heavy blooms on July 13-14 and on August 18-19, 2011, at the deepest part of the lake, MDT, to describe the diel bloom cycle and dissolved nutrient concentrations. Sondes (YSI, Inc.) deployed at sites MDN and MDT recorded pH, specific conductance, temperature, and dissolved oxygen saturation hourly at two depths, 1 m from the surface and 1 m from the bottom. An additional sensor in the shallow sonde recorded phycocyanin (a photosynthetic pigment specific to cyanobacteria) fluorescence hourly. Phycocyanin depth profiles were also completed weekly during water sample collection.

### Regression Models

Regression models for each dependent variable were developed separately from data collected at each site using the *step* function in the *stats* package of R version 2.12.1, which adds terms to and removes terms from the generalized linear model of the regression in a systematic fashion and chooses the best model based on a minimization of the Akaike Information Criterion. The starting set of independent data for each model included pH, temperature, dissolved oxygen saturation, discrete-depth phycocyanin fluorescence (at 1 m), and sin and cos functions of the Julian day. Models were developed by considering the independent data as a set of 1) instantaneous values on the hour closest to the sample collection time, 2) median values of a 6-hour window surrounding the sample collection time, and 3) daily median values on each sample date. The data set (instantaneous, 6-hour medians, or daily medians) that resulted in the best model was retained, and the explanatory variables were further scrutinized for multi-collinearity by examining variance inflation factors and the correlation matrix. Parsimony was achieved by dropping any explanatory variables retained by the step function that were not significant at  $p < 0.05$  and for which the variance inflation factor was greater than five. Regression variables were also visually inspected for serial autocorrelation. One variable was removed from the model when correlation coefficients between two explanatory variables exceeded 0.5 (absolute value). Regression models were constructed using provisional data (do not cite).

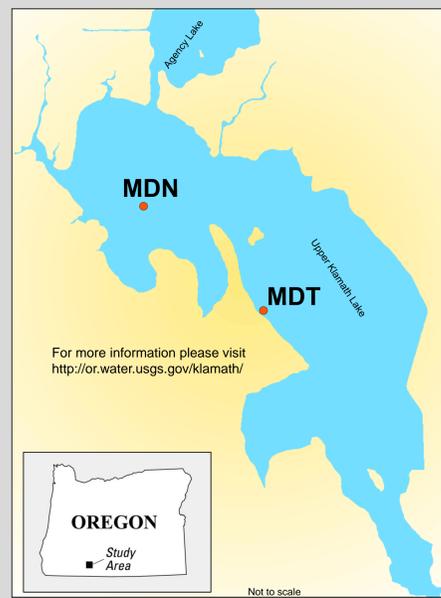


Figure 1. Location of water-quality monitoring sites in Upper Klamath Lake, OR, 2011

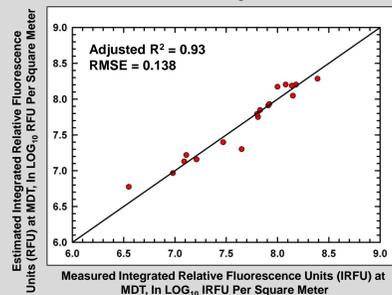
## Results

Regression models used to estimate integrated phycocyanin fluorescence and depth-averaged concentrations of chlorophyll *a* and total phosphorus explained their variability relatively well, in that adjusted R<sup>2</sup> values ranged between 0.94 and 0.79 (fig. 2). Models based on site MDT data had slightly higher adjusted R<sup>2</sup> values than those based on data from MDN. Models of depth-integrated fluorescence had adjusted R<sup>2</sup> values of 0.93 and 0.89 for MDT and MDN, respectively, and models of depth-averaged chlorophyll *a* concentration had adjusted R<sup>2</sup> values of 0.94 and 0.91 respectively. Adjusted R<sup>2</sup> values of total phosphorus models were lower at 0.84 and 0.79 for MDT and MDN respectively. The regression model for total phosphorus at MDN in particular, is not satisfactory because there is evidence of serial autocorrelation in the residuals (fig. 3A). All models included pH and/or dissolved oxygen saturation as independent variables, which are closely related to bloom growth, and sin and/or cos terms, which provided a temporal cycle. Only one of the two integrated phycocyanin fluorescence models included the discrete-depth fluorescence data (fig. 2D).

Time series shown at a daily step of estimated integrated phycocyanin fluorescence, chlorophyll *a* concentration, and total phosphorus concentration (fig. 3) show two distinct bloom periods in 2011 and short-term variability that was not captured by weekly sampling. Samples for integrated phycocyanin fluorescence and depth-averaged chlorophyll *a* concentration collected over the 24-hr intervals at MDT fell largely within the 90 percent confidence intervals, indicating that, although the models were based on samples collected weekly between approximately 9:30AM and 10:30AM at MDN and between 11:30PM and 12:30PM at MDT, they have some predictive power for environmental conditions other times of the day.

### Regression Estimated vs. Measured Values

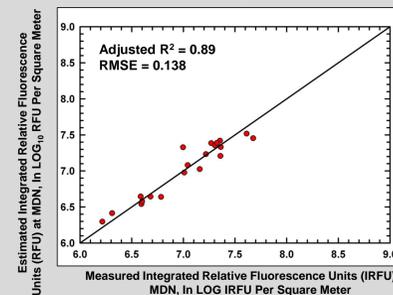
(A) MDT Regression Estimated vs. Measured Integrated RFU



$$\text{Log}_{10} \text{IRFU} = -0.700 \sin(2\pi(D/365)) - 0.486 \cos(2\pi(D/365)) + 0.447 \text{pH}_{1,d} + 3.36$$

Where:  $\text{pH}_{1,d}$  is the daily median for pH from the lower sonde at MDT, and D is the Julian date

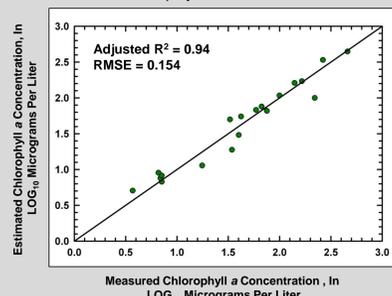
(D) MDN Regression Estimated vs. Measured Integrated RFU



$$\text{Log}_{10} \text{IRFU} = -0.317 \sin(2\pi(D/365)) - 0.204 \cos(2\pi(D/365)) + 0.043 \text{RFU}_{1,d} + 0.340 \text{pH}_{1,d} + 3.69$$

Where:  $\text{RFU}_{1,d}$  is the daily median for RFU from the upper sonde at MDN,  $\text{pH}_{1,d}$  is the daily median for pH from the lower sonde at MDN, and D is the Julian date

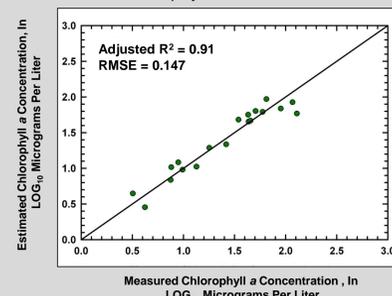
(B) MDT Regression Estimated vs. Measured Chlorophyll a Concentration



$$\text{Log}_{10} \text{Chlorophyll } a = -0.419 \sin(2\pi(D/365)) + 0.024 \text{RFU}_{1,d} + 0.396 \text{pH}_{1,d} - 2.10$$

Where:  $\text{RFU}_{1,d}$  is the daily median for RFU from the upper sonde at MDT,  $\text{pH}_{1,d}$  is the daily median for pH from the lower sonde at MDT, and D is the Julian date

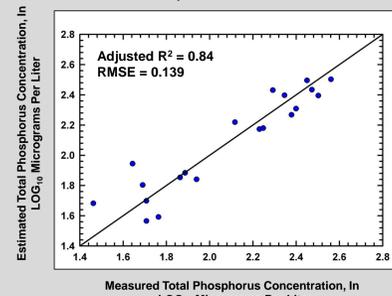
(E) MDN Regression Estimated vs. Measured Chlorophyll a Concentration



$$\text{Log}_{10} \text{Chlorophyll } a = -0.711 \sin(2\pi(D/365)) - 0.625 \cos(2\pi(D/365)) + 0.400 \text{pH}_{6,d} - 2.66$$

Where:  $\text{pH}_{6,d}$  is the six-hour median for pH from the lower sonde at MDN, and D is the Julian date

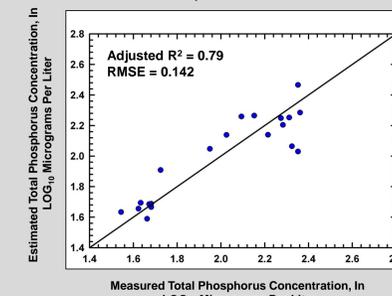
(C) MDT Regression Estimated vs. Measured Total Phosphorus Concentration



$$\text{Log}_{10} \text{TP} = 0.450 \cos(2\pi(D/365)) + 0.034 \text{T}_{1,d} - 0.006 \text{Sat}_{1,d} + 0.018$$

Where:  $\text{T}_{1,d}$  is the six-hour median for temperature from the upper sonde at MDT,  $\text{Sat}_{1,d}$  is the six-hour median for oxygen saturation from the upper sonde at MDT, and D is the Julian date

(F) MDN Regression Estimated vs. Measured Total Phosphorus Concentration



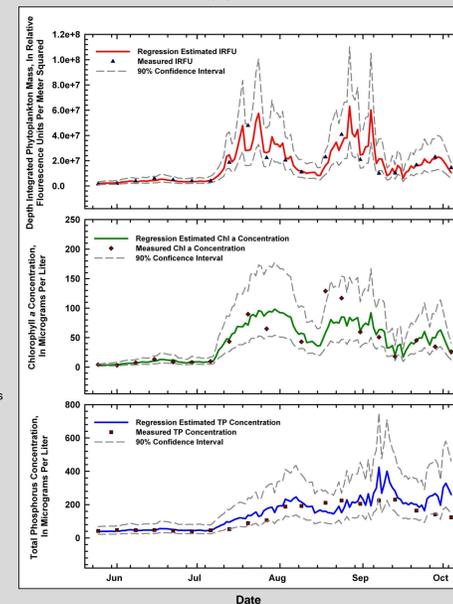
$$\text{Log}_{10} \text{TP} = 0.344 \cos(2\pi(D/365)) + 0.311 \text{pH}_{1,d} + 0.131 \text{Sat}_{1,d} + 0.240$$

Where:  $\text{pH}_{1,d}$  is the daily median for pH from the upper sonde at MDN,  $\text{Sat}_{1,d}$  is the daily median for dissolved oxygen saturation from the lower sonde at MDN, and D is the Julian date

Figure 2. Regression estimated values compared to measured values for depth-integrated cyanobacteria cell density, depth-averaged chlorophyll *a* concentration, and depth-averaged total phosphorus concentration for MDT (A, B, and C) and MDN (D, E, and F).

### Season Reconstruction Using Regression Estimated Values

(A) MDN



(B) MDT

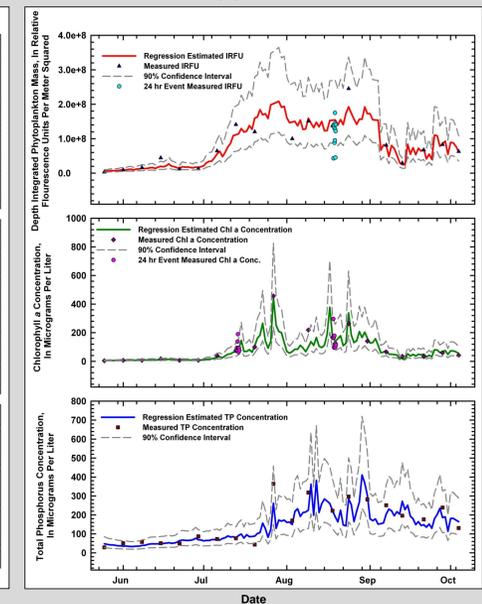


Figure 3. Time series of regression estimated cell density values, chlorophyll *a*, total phosphorus, compared to measured values, and 90 percent confidence intervals for MDN (A) and MDT (B).

## Conclusions

- Adjusted R<sup>2</sup> values and comparison of estimated values to measured values show that in-situ continuous sonde data can be used to estimate depth integrated cell density, depth-averaged chlorophyll *a* concentration, and depth-averaged total phosphorus concentration
- The estimated time series of biomass measurements based on in-situ monitor data show variability on time scales of one to five days that can't be captured with analysis of weekly samples and, therefore provide a cost effective way obtaining more time-dense information
- Field measurements of chlorophyll *a* and integrated phycocyanin fluorescence per square meter have high diel variability, however, samples collected at other time s of the day were, for the most part, within a 90 percent confidence interval around the estimated values based on a model developed with samples typically collected mid-morning
- Phycocyanin fluorescence data may be limited for use as a cyanobacteria biomass surrogate, but the use of these sensors may be the first step toward surrogate development.
- Improvements to the models used here include:
  - More years of data to test for model robustness
  - Stratification (differences in pH, temperature, and oxygen saturation between deep and shallow depths) is used as an independent variable

## Acknowledgements

Continued funding and support is provided by the Bureau of Reclamation, Klamath Basin Area Office. Assistance from Eric Janney and the personnel of the USGS Klamath Falls Field Station is gratefully acknowledged for facilitating the water-quality field program. Kristen Harbin also contributed greatly to the project.

## References

- Banish, N.P., Adams, B.J., Shively, R.S., Mazur, M.M., Beauchamp, D.A., and Wood, T.M., 2009. Distribution and habitat associations of radio-tagged adult Lost River and Shortnose Suckers in Upper Klamath Lake, Oregon: Bethesda, Maryland, Transactions of the American Fisheries Society, v. 138.
- Christensen, V.G., Jian, Xiaodong, Ziegler, A.C., 2000. Regression Analysis and Real-Time Water-Quality Monitoring to Estimate Constituent Concentrations, Loads, and Yields in the Little Arkansas River, South-Central Kansas, 1995-99. U.S. Geological Survey Water-Resources Investigations Report 00-4126, 36 pp
- Kann, Jacob, and Walker, W., 2001. Nutrient and hydrological loading of Upper Klamath Lake, Oregon, 1991-1998, prepared for the Klamath Tribes Resources Department and the Bureau of Reclamation: Klamath Falls, Oregon, 48 p. plus appendices.
- Perkins, D.L., Kann, J., and Scopettone, G.G., 2000. The role of poor water quality and fish kills in the decline of endangered Lost River and shortnose suckers in Upper Klamath Lake: Klamath Falls, Oregon, U.S. Geological Survey, Biological Resources Division Report Submitted to Bureau of Reclamation, Klamath Falls Project Office, contract 4-AA-29-12160, 39 p.
- Ryberg, K.R., 2006. Continuous water-quality monitoring and regression analysis to estimate constituent concentrations and loads in the Red River of the North, Fargo, North Dakota, 2003-05. U.S. Geological Survey Scientific Investigations Report 2006-5241, 35 p.
- Walker, W.W., 2001. Development of a phosphorus TMDL for Upper Klamath Lake, Oregon: Bend, Oregon, prepared for Oregon Department of Environmental Quality, accessed April 23, 2012 at <http://www.deq.state.or.us/wq/TMDLs/docs/klamathbasin/ukl drainage/devphostmdl.pdf>
- Wood, T.M., and Gartner, J.W., 2010. Use of acoustic backscatter and vertical velocity to estimate concentration and dynamics of suspended solids in Upper Klamath Lake, south-central Oregon: Implications for *Aphanizomenon flos-aquae*. U.S. Geological Survey Scientific Investigations Report 2010-5203, 20p.