LIMNOLOGY
and
OCEANOGRAPHY: METHODS

ASLO

Limnol. Oceanogr.: Methods 2018

doi: 10.1002/lom3.10287

Using in situ UV-Visible spectrophotometer sensors to quantify riverine phosphorus partitioning and concentration at a high frequency

Matthew C. H. Vaughan [1,2]* **William B. Bowden** [1] **James B. Shanley** [3] **Andrew Vermilyea** [4] **Beverley Wemple** [1] **Andrew W. Schroth** [5]

[1]University of Vermont, Rubenstein School of Environment and Natural Resources, Burlington, Vermont
[2]Lake Champlain Basin Program, Grand Isle, Vermont
[3]US Geological Survey, Montpelier, Vermont
[4]Castleton University, Department of Natural Sciences, Castleton, Vermont
[5]University of Vermont, Department of Geology, Burlington, Vermont

Abstract

Accurate riverine phosphorus concentration measurements are often critical to meet watershed management goals. Phosphorus monitoring programs often rely on proxy variables such as turbidity and discharge and have limited ability to accurately estimate concentrations of dissolved phosphorus fractions that are most bioavailable. Optical water quality sensors can make subhourly measurements and have been shown to reduce uncertainty in load estimates and reveal high-frequency storm dynamics for nitrate and dissolved organic carbon. We evaluated the utility of in situ UV-Visible spectrophotometers to predict total, dissolved, and soluble reactive phosphorus concentrations in streams draining agricultural, urban, and forested land use/land covers. We present the first statistically validated application of optical water quality sensors to demonstrate how sensors may perform in predicting phosphorus fraction concentrations through training set models. Total phosphorus predictions from UV-Visible spectra were optimal when models were site-specific, and the proportion of variance explained was generally as high as or higher than the results of other studies that rely only on discharge and turbidity. However, root mean square errors for total phosphorus models were relatively high compared to the median concentrations at each site. Models to predict dissolved and soluble reactive phosphorus concentrations explained a greater proportion of the variance than any other known proxy variable technique, and results varied by land use/land cover. Though accuracy limitations remain, this approach has potential to predict concurrent total, dissolved, and soluble reactive phosphorus concentrations at a high frequency for many applications in water quality research and management communities.

Elevated phosphorus concentrations cause persistent problems such as eutrophication and potentially toxic cyanobacteria growth in many fresh waterbodies that impact recreation, drinking water quality, property values, and ecosystem health (Carpenter et al. 1998; Conley et al. 2009). To address these challenges, watersheds are often managed to reduce tributary total phosphorus (TP) loads (Sharpley et al. 1994; Djodjic et al. 2002). Accurate tributary phosphorus load estimation is critical to meet these management goals, and TP load estimates assess the efficacy of watershed-scale phosphorus reduction efforts (Medalie 2016). Episodic storm events are particularly important to capture, since they deliver disproportionately large loads of water, sediment, and phosphorus (Jordan et al. 2007; Sharpley et al. 2008) and phosphorus concentrations change rapidly during storms (Correll et al. 1999).

TP is delivered to waterbodies in several forms that can differ in bioavailability for cyanobacteria growth (Correll 1998; Giles et al. 2015; Isles et al. 2017). Phosphorus is most bioavailable as dissolved inorganic orthophosphate (PO_4^{3-}), commonly measured as soluble reactive phosphorus (SRP), or as part of the total dissolved phosphorus (TDP) fraction. A portion of the organic phosphorus pool can also be directly bioavailable, or can be rapidly decomposed by heterotrophic bacteria into the inorganic form that can be quickly utilized (Kane et al. 2014). Particulate phosphorus has potential bioavailability dependent upon the speciation of solid phase phosphorus and its interaction with receiving water column and pore-water solutions (Giles et al. 2015; Schroth et al. 2015). Because each phosphorus fraction has differing

*Correspondence: mvaughan@lcbp.org

Additional Supporting Information may be found in the online version of this article.

degrees of bioavailability, understanding the magnitude and dynamic chemical partitioning of riverine phosphorus fraction loads delivered to a receiving waterbody is necessary to inform management of potential cyanobacteria growth and to reach desired management outcomes (Stumpf et al. 2012; Isles et al. 2017). Long-term continuous monitoring is particularly important to characterize changes as management decisions and land-use change influence the amount and composition of phosphorus delivery to receiving waterbodies (Dodd and Sharpley 2016; Jarvie et al. 2017).

TP concentration estimates are often based on correlations of lab-measured TP concentration from grab samples with continuously measured discharge, turbidity, or a combination of the two. Although these correlations can be strong (Hyer et al. 2016), this approach has two disadvantages: (1) solute-discharge and solute-turbidity relationships are variable among storm events due to hysteresis effects and threshold behavior changes (Dhillon and Inamdar 2013; Bieroza and Heathwaite 2015) and (2) these methods only estimate TP concentrations and typically do not provide critical information on phosphorus partitioning. Alternatively, SRP concentration can be directly measured in situ at an hourly to subhourly frequency with newly available wet chemistry instruments (e.g., Cohen et al. 2013).

In situ spectrophotometer sensors offer the potential to concurrently measure multiple phosphorus fraction concentrations (e.g., TP, TDP, and SRP) at a high frequency continuously with no reagents or waste products. These sensors measure light absorbance in the UV-Visible spectrum and have been shown to make continuous, concurrent, and accurate measurements of dissolved organic carbon, nitrate (NO_3^-), and total suspended solids concentrations in surface waters with varying environmental conditions and aqueous matrices (Langergraber et al. 2003; Rieger et al. 2006; Sakamoto et al. 2009; Fichot and Benner 2011). Because optical sensors can be deployed on a long-term basis and operate continuously, researchers and watershed managers can better characterize large episodic events when manual sampling may be impractical, expensive, and/or unsafe (Saraceno et al. 2009; Carey et al. 2014). Dynamics that occur on seasonal or diel timescales are also better described by this approach (Heffernan and Cohen 2010; Pellerin et al. 2012). While methods that rely on discrete grab samples may assign a single concentration to an entire storm or day of record, high-frequency measurements capture short timescale hysteresis and threshold behavior changes not documented by discrete samples. In addition, optical sensors have the potential to predict nutrient concentrations and vertical profiles in lakes and reservoirs, where concentration-discharge relationships are not applicable (Birgand et al. 2016; Joung et al. 2017). Continuous and high-frequency monitoring can improve accuracy of load estimates (Guo et al. 2002; Pellerin et al. 2014), though the improvement over concentration-discharge measurement may be limited for some applications (Musolff et al. 2017).

Only a few researchers have attempted to use multiwavelength UV-Visible spectrophotometers to estimate phosphorus fraction concentrations in a limited number of environmental conditions, and it is unknown how performance may differ among streams draining different land uses and land covers (LULCs). Unlike solutes such as nitrate and dissolved organic carbon, most phosphorus fractions do not directly absorb light in the UV-Visible spectrum, so calibrations with concentrations of different phosphorus fractions rely on proxy correlations alone, similar to correlations relating TP concentration to discharge. This approach has also been used to predict other non-UV-Visible wavelength light absorbing solutes (e.g., Si, Mn, and Fe) with promising results (Birgand et al. 2016). Because spectrophotometers measure absorbance throughout the entire UV-Visible spectrum, it is possible that multiple light sensitive proxies covary with phosphorus fractions differently by site, season, and/or storm event. Different phosphorus fractions may be tracking light sensitive aqueous components that reflect phosphorus provenance and biogeochemical cycling within a particular catchment and across different temporal scales or flow regimes. UV-Visible spectrophotometer sensors have shown promise to predict phosphorus fractions in some cases (Etheridge et al. 2014), though predictions of TP, TDP, and SRP concentrations from optical sensors have not been evaluated rigorously in a variety of systems. It is not known to what extent site-specific calibrations are necessary as is often the case for other solutes (e.g., Vaughan et al. 2017), or whether multiple different phosphorus fraction concentrations can be predicted accurately from UV-Visible absorbance spectra. If robust proxy correlations were developed, phosphorus fractions could be measured continuously on short timescales that capture rapid changes in hydrologic and biogeochemical processes critical to inform watershed management and nutrient reduction goals.

Generating algorithms to predict nutrient concentrations from absorbance spectra presents a challenge due to the high dimensionality of the independent variables (light absorbance spectra) compared to the single response variable (nutrient concentration). Partial least squares regression (PLSR) can be used to harness the information of a rich collection of independent variables to predict a desired dependent quantity. PLSR is a technique that condenses independent variables into orthogonal, uncorrelated components and combines them in a multivariate model to predict the parameter of interest. Visible, near-infrared, and far-infrared reflectance spectra have been used extensively in combination with the PLSR approach to describe soil characteristics such as available phosphorus, electrical conductivity, pH, organic carbon, lime requirement, and cation exchange capacity (e.g., McCarty et al. 2002; Viscarra Rossel et al. 2006). In addition, UV-Visible spectra have been used to predict concentrations of various nutrients in fresh and brackish water with encouraging results (Avagyan et al. 2014; Birgand et al. 2016; Vaughan et al. 2017). However, previous studies evaluating this method to predict

constituent concentrations in water have not presented model validation results; that is, all of the available laboratory analyses were used to calibrate the model, without verifying predictions using independent observations.

We deployed spectrophotometers in well-characterized watersheds of different LULCs (Rosenberg and Schroth 2017; Vaughan et al. 2017) that drive different phosphorus dynamics, concentrations, and partitioning. UV-Visible absorption spectra from spectrophotometers were coupled with grab water samples for conventional laboratory analysis of TP, TDP, and SRP concentrations. Our objectives were to: (1) evaluate in situ UV-Visible spectrophotometer prediction of TP, TDP, and SRP concentrations in riverine waters, (2) compare prediction performance in surface waters draining watersheds of various LULCs, and (3) present statistical validation of these predictions. To our knowledge, this work constitutes the most rigorous assessment to date of the utility of this sensor technology to predict multiple phosphorus fraction concentrations across a range of riverine environments.

Study areas

The study sites were in the Lake Champlain Basin of Vermont in the northeastern US (Table 1, Fig. 1). The study streams were selected because their watershed LULC was dominantly agricultural, urban, or forested, and each watershed met criteria for watershed size, accessibility, and discharge data availability. Hungerford Brook is a primarily agricultural catchment, including dairy production, row crops, hay, and pasture. Potash Brook is situated near the city of Burlington, which is Vermont's densest population center. Its watershed is primarily characterized by urban and suburban development (54%), and includes some agricultural and forest cover (29% and 11%, respectively). The Wade Brook catchment is primarily forested (95%) and is situated on the western slope of Vermont's Green Mountain chain.

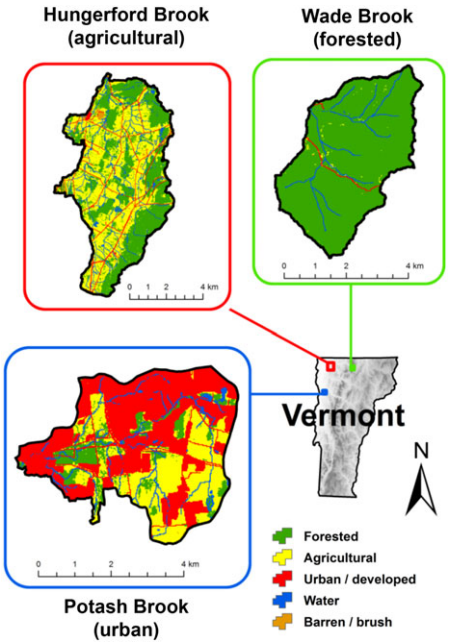

Fig. 1. Map showing location and land use/land cover of the three study areas.

Table 1. Summary of study area characteristics.

	Hungerford Brook	Potash Brook	Wade Brook
Primary land cover	Agricultural	Urban/suburban	Forested
Watershed area (km^2)	48.1	18.4	16.7
Percent forested	40.5	10.6	95.1
Percent agricultural	44.8	29.1	0.6
Percent urban	5.6	53.5	0.8
Percent impervious area	2.3	23.9	0.0
Sensor elevation (m)	80	42	320
Maximum watershed elevation (m)	354	143	981
Mean watershed slope (%)	5.6	5.3	26
Mean air temperature (°C)	6.7	7.8	4.2
Mean annual precipitation (mm)	1000	961	1453
Sensor optical path length (mm)	5.0	5.0	15.0
Coordinates (WGS 1984)	44.918403°N, 73.055664°W	44.444331°N, 73.214482°W	44.864468°N, 72.552904°W
Soil and surficial geology	Sandy, silty, and stony loams	Sandy and silty loams, clay	Glacial till, sandy loam
Vegetation	Agricultural, mixed northern hardwoods and conifer	Urban/suburban landscaping, mixed northern hardwoods and conifer, agricultural	Mixed northern hardwoods and conifer

Hungerford Brook and Wade Brook drain to the Missisquoi River and Lake Champlain; Potash Brook drains directly to Lake Champlain. Precipitation totals in the Wade Brook catchment are higher than the catchments of Hungerford Brook and Potash Brook due to orographic effects (Table 1).

Methods

In-stream measurements

We used s::can Spectrolyser UV-Visible spectrophotometers (s::can Messtechnik GmbH, Vienna, Austria) in each stream, deployed from June 2014 to December 2016 for spring, summer, and fall seasons. The sensors were housed in PVC tubing for protection during high flows, were solar powered for autonomous operation, and transmitted summary data through a cellular data network. Full UV-Visible spectra measurements were stored on-board the sensor and downloaded manually on site. The spectrophotometers measured light absorbance at wavelengths ranging from 220 nm to 750 nm at 2.5 nm increments and were programmed to take measurements every 15 min. Optical path lengths were either 5 mm or 15 mm, depending on the typical turbidity of each stream (Table 1), and absorbance spectra were normalized by optical path length for comparison. Sensor measurement windows were automatically cleaned before each measurement with a silicone wiper and cleaned manually in the field at least every 2 weeks using pure ethanol. To focus on dissolved constituents, raw absorbance spectra were corrected for the effects of turbidity by fitting a third-order polynomial in the visible range of the spectrum, extrapolating into the UV portion, and then subtracting the extrapolated absorbance from the raw spectrum (Langergraber et al. 2003; Avagyan et al. 2014).

Laboratory measurements

Manual grab samples were collected at the sensor sites across the monitored seasons during baseflow and storms (peak flow, rising, and falling limb), timed to coincide with sensor measurements to calibrate in situ UV-Visible absorbance spectra to laboratory TP, TDP, and SRP concentration measurements (Fig. 2). Care was taken to collect samples directly adjacent to the sensor measurement window. We analyzed a total of 560 grab samples over the course of the study. Samples taken in 2015 were analyzed for TDP and SRP; samples taken in 2016 were analyzed for TP, TDP, and SRP. We filtered TDP and SRP samples in the field using sample-rinsed glass fiber GF/F filters (nominal pore size of 0.7 μm) into new, triple-rinsed HDPE bottles, and collected TP samples from the stream without filtering. This filter size differs from that of some others studies where 0.45 μm filters are used. This may influence absolute lab value comparability (where values in this study may be slightly higher in comparison), but would not influence the evaluation of model calibration or validation techniques, which is the focus of this work. We stored samples on ice in the field and in transport, then stored either in a cooler at 2°C (for TDP and SRP samples) or in a freezer at −23°C (for TP samples) until analysis.

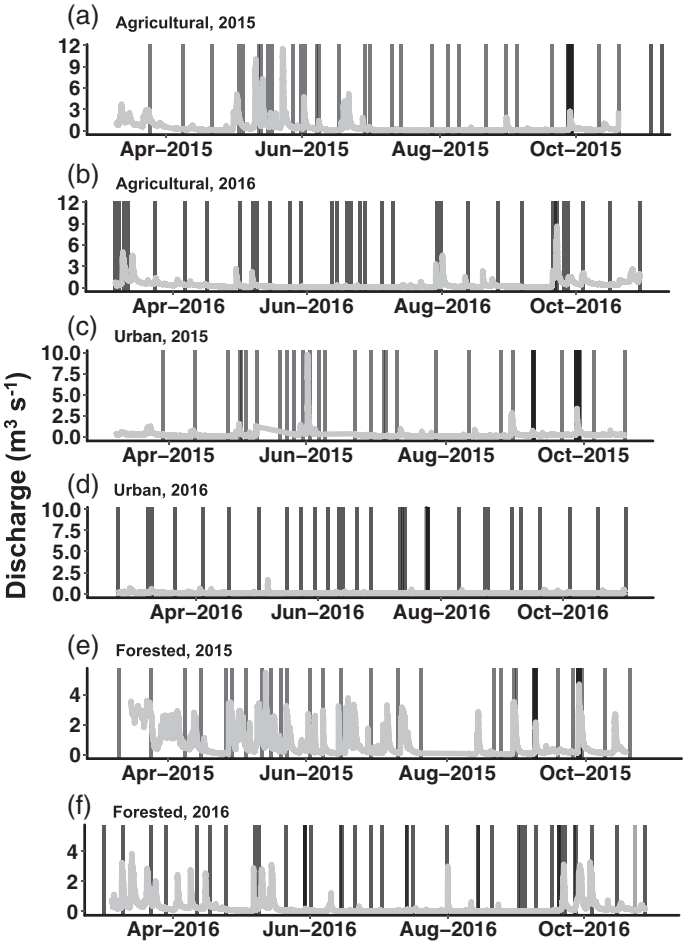

Fig. 2. Discharge (gray lines) and manual grab sample times (black vertical lines) at the (a-b) agricultural, (c-d) urban, and (e-f) forested sites. Samples taken in 2015 were analyzed for TDP and SRP; samples taken in 2016 were analyzed for TP, TDP, and SRP.

We analyzed for TP concentration by first liberating organic phosphorus as inorganic phosphorus through oxidation by persulfate, followed by the molybdate method (US EPA method 365.1 4500-PJ). We measured TDP concentration the same way as TP after samples had been filtered as described above. SRP concentration was determined colorimetrically by measuring absorbance of 885 nm following sample reaction with molybdate, ascorbic acid, and trivalent antimony, also consistent with US EPA method 365.1 (Parsons et al. 1984). For each analyte, the nonparametric Kruskal–Wallis test (Kruskal and Wallis 1952) was used to determine whether concentrations were significantly different among the three sites.

Phosphorus fraction concentration prediction: Training and validation techniques

When reporting correlations of a particular method to predict lab measurements, it is common to develop a model using

all available data and then assume model statistics will apply to future predictions using unknown data. In contrast, we used a bootstrapping technique to validate the accuracy of calibration models built on only a portion of the data to provide a more robust method to assess uncertainty in concentration prediction. Training and validation prediction sets were generated for TP, TDP, and SRP using combined data from all sites for each parameter, and by separating the available data by site. For each training dataset, 85% of available observations were selected randomly to generate a model. The model was developed by an identical approach to Etheridge et al. (2014), where PLSR was employed with the pls package in R to generate calibration algorithms (Mevik et al. 2015; R Core Team 2015). Each model incorporated a number of components equal to a maximum of approximately 10% of the observations as recommended by Mevik et al. (2015).

The training model was then used to predict a validation set, which was comprised of the remaining 15% of observations that were randomly withheld. This process was repeated 1000 times with replacement for each parameter, and predictions and statistics for each model were collected and aggregated. We then calculated the means and standard deviations of predicted concentration values for all 1000 iterations of training and validation sets. Sensor performance was evaluated by performing linear correlations on the mean predicted value for training and validation sets vs. the corresponding lab-measured values. Throughout the article, adjusted R^2 values are presented to compare goodness of fit for regressions to remove the bias associated with differing sample sizes (Ohtani 2000), and root mean square errors (RMSEs) are presented as estimates of model accuracy. The result of this process is a quantifiable level of confidence for how accurately UV-Visible absorbance spectra may predict TP, TDP, and SRP concentrations at times when no lab measurement is available for comparison. This type of model validation is common in many disciplines and is more robust than other approaches that develop models using the entire available dataset and therefore provide stronger prediction statistics (Aber 1997).

Logarithmically transformed discharge or turbidity measurements are often used to predict riverine TP concentrations (Hirsch et al. 2010; Stutter et al. 2017). We performed multiple linear regressions using these two variables to predict TP concentrations at each site and compared this method with the performance of the UV-Visible spectrophotometers. These models included all available data for each site in order to form comparisons using the most favorable case for this method.

Table 2. Summary statistics for grab samples collected at the study sites (all concentrations are in μg P L^{-1}).

	Agricultural	Urban	Forested
Total phosphorus (2016 only)			
Count	36	27	42
Minimum	13.4	6.50	0.70
Maximum	917	89.6	12.3
Median	79.0	20.6	3.8
Mean	130	24.7	4.5
Variance	31.6	0.40	< 0.10
Total dissolved phosphorus (2015–2016)			
Count	77	80	89
Minimum	8.50	3.5	1.8
Maximum	1413	263	31.6
Median	63.0	32.1	7.6
Mean	133	64.6	10.7
Variance	42.2	5.4	0.10
Soluble reactive phosphorus (2015–2016)			
Count	77	105	89
Minimum	3.0	0.30	0.60
Maximum	1240	231.5	22.8
Median	46.2	17.4	4.2
Mean	110	37.1	6.9
Variance	35.3	2.6	< 0.10

Results

Phosphorus grab samples and UV-Visible absorbance measurements

The nonparametric Kruskal–Wallis test revealed that grab sample concentrations for TP, TDP, and SRP were each significantly different among these three sites ($p < 0.001$; Table 2). When UV-Visible absorbance spectra were plotted and colored by corresponding phosphorus fraction concentrations, it is evident that much of the variability in UV-Visible spectra occurs in the wavelength range of 220–350 nm (Fig. 3). Furthermore, while generally higher absorbances correspond with higher phosphorus fraction concentrations, complex relationships exist between spectral data and phosphorus fraction concentrations. The ratio of lab-measured TDP to TP concentrations and the ratio of lab-measured SRP to TP concentrations varied considerably at the agricultural and urban sites, and the ratio of SRP to TP was significantly different between these sites as determined by the nonparametric Mann–Whitney U test ($p = 0.001$) (Fig. 4). The highest observed concentrations of TDP and SRP were higher than the highest TP concentration because relatively large storms in 2015 produced high TDP and SRP concentrations and TP concentrations were not measured at that time.

Total phosphorus

Predictive models for TP were developed using data for all sites, with 10 components (11% of observations). The training sets explained a relatively high proportion of the variance in TP concentration (adj. $R^2 = 0.96$; $p < 0.001$), while correlations from the bootstrap validation method explained approximately three-quarters of the variance in TP concentration (adj. $R^2 = 0.78$; $p < 0.001$) (Fig. 5). RMSEs were 25 μg P L^{-1} for training sets and 59 μg P L^{-1} for validation sets. The RMSE of the validation set was 75% of the median TP concentration

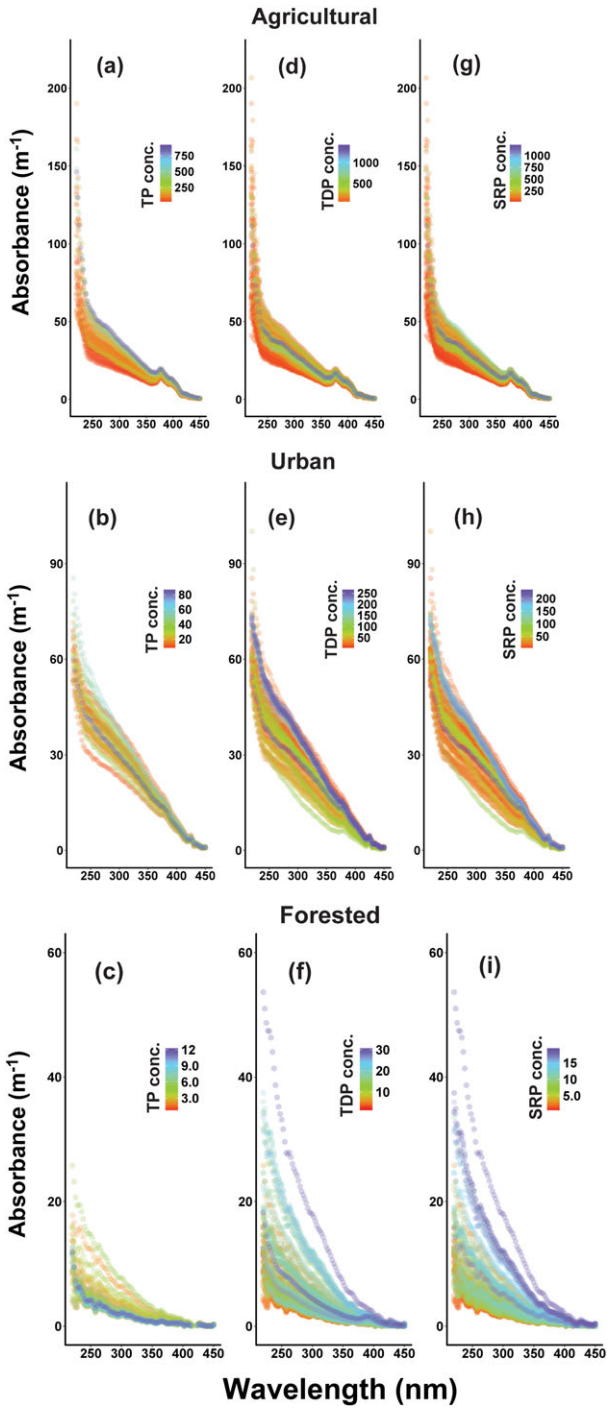

Fig. 3. Plots of compensated UV-Visible absorbance spectra vs. wavelength of light and corresponding (**a**–**c**) TP, (**d**–**f**) TDP, and (**g**–**i**) SRP concentrations (μg P L^{-1}) in color for agricultural, urban, and forested sites.

for the agricultural site, and was 3 and 16 times greater than the median TP concentrations at the urban and forested sites, respectively. Separating datasets by site did not improve goodness of fit for predictive models, though it resulted in validation RMSE values that were 74–80% of the median TP concentrations at the urban and forested sites (Table 3).

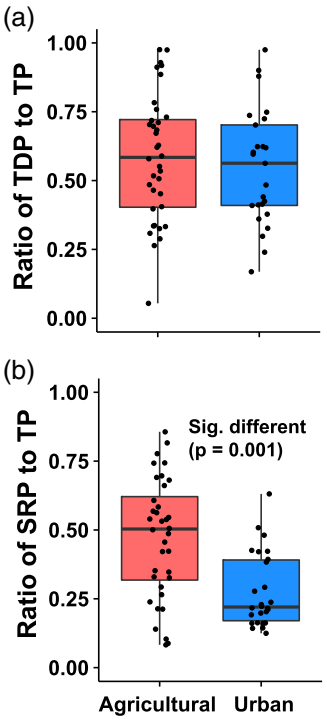

Fig. 4. Box and whisker plots of the ratios of (**a**) TDP to TP and (**b**) SRP to TP for the agricultural and urban sites in 2016. Data from the forested site is not shown, since concentration differences between different operationally defined fractions were within the range of analytical error.

When logarithmically transformed discharge and turbidity measurements were used to predict TP concentration using a multiple linear regression model for each site separately, the adjusted coefficients of determination were 0.05, 0.14, and 0.41 for the forested, urban, and agricultural sites, respectively. However, models for the forested and urban sites were not statistically significant and the accuracy for the agricultural site model was lower than for models based on the UV-Visible absorbance spectra (RMSE = 138 μg P L^{-1}).

Total dissolved phosphorus

Predictive models for TDP using data for all sites were developed with 18 components (8% of observations). The training sets explained a relatively high proportion of the variance in TDP concentrations (adj. $R^2 = 0.96$; $p < 0.001$), while correlations from the bootstrap validation method explained nearly two-thirds of the variance in TDP concentration (adj. $R^2 = 0.61$; $p < 0.001$) (Table 3; Fig. 6a,b). Separating datasets by site increased accuracy and the proportion of the variance explained in validation sets for the urban site (adj. $R^2 = 0.68$; $p < 0.001$) and the forested site (adj. $R^2 = 0.74$; $p < 0.001$) with eight components, but did not improve validation performance at the agriculture site. Accuracies for these models were limited, however. Validation set RMSEs were greater than the median TDP concentrations for the agricultural and urban

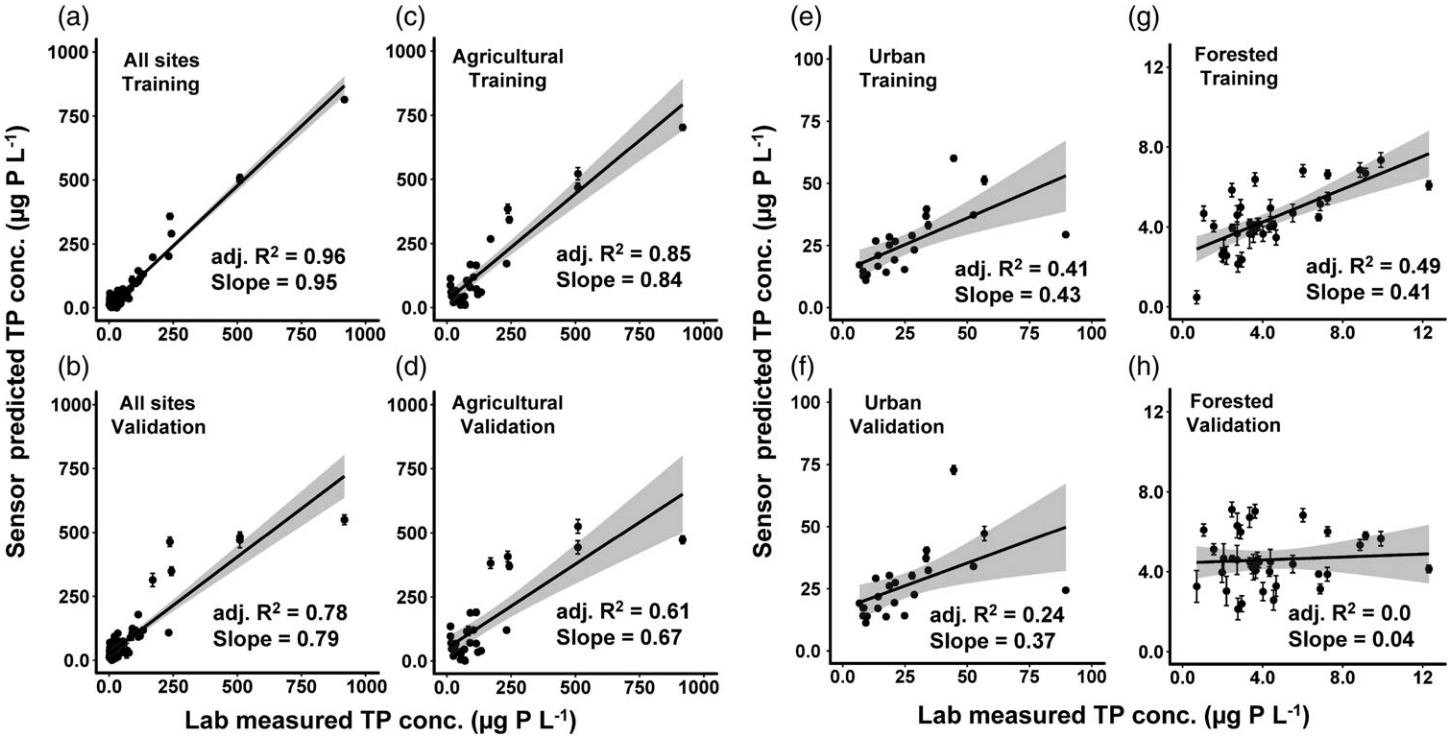

Fig. 5. Bootstrap TP training and validation plots for (**a**, **b**) all combined, (**c**, **d**) agricultural, (**e**, **f**) urban, and (**g**, **h**) forested sites. Correlations were statistically significant ($p < 0.001$) for all but the forested validation sets (**h**). Shading represents 90% confidence intervals. Error bars represent one standard deviation for the predictions over 1000 bootstrap iterations. Note that error bars are present for all points but may not be visible and that scales differ among plots.

sites, and was 55% of the median TDP concentration for the forested site. Plotting residuals in the TDP models by lab-measured value, turbidity, and discharge did not reveal discernable patterns in prediction error (Supporting Information Fig. B).

Soluble reactive phosphorus

Predictive models for SRP concentration using data for all sites were developed with 18 components (7% of observations). The training sets explained a relatively high proportion of the variance in SRP concentration (adj. $R^2 = 0.96$; $p < 0.001$), while

Table 3. Summary of PLSR model results.

Site(s)	Observations	Components	Training adj. R^2	Training RMSE (μg P L^{-1})	Validation adj. R^2	Validation RMSE (μg P L^{-1})
Total phosphorus						
All	90	10	0.96	25	0.78	59
Agricultural	31	4	0.85	70	0.61	115
Urban	24	3	0.41	14	0.24	17
Forested	36	4	0.49	1.9	−0.02	2.8
Total dissolved phosphorus						
All	222	18	0.96	29	0.61	90
Agricultural	70	8	0.88	73	0.56	147
Urban	73	8	0.90	24	0.68	43
Forested	79	9	0.94	1.8	0.72	4.2
Soluble reactive phosphorus						
All	247	18	0.96	23	0.68	68
Agricultural	70	8	0.92	54	0.70	109
Urban	98	10	0.94	13	0.57	36
Forested	79	9	0.95	1.2	0.79	2.4

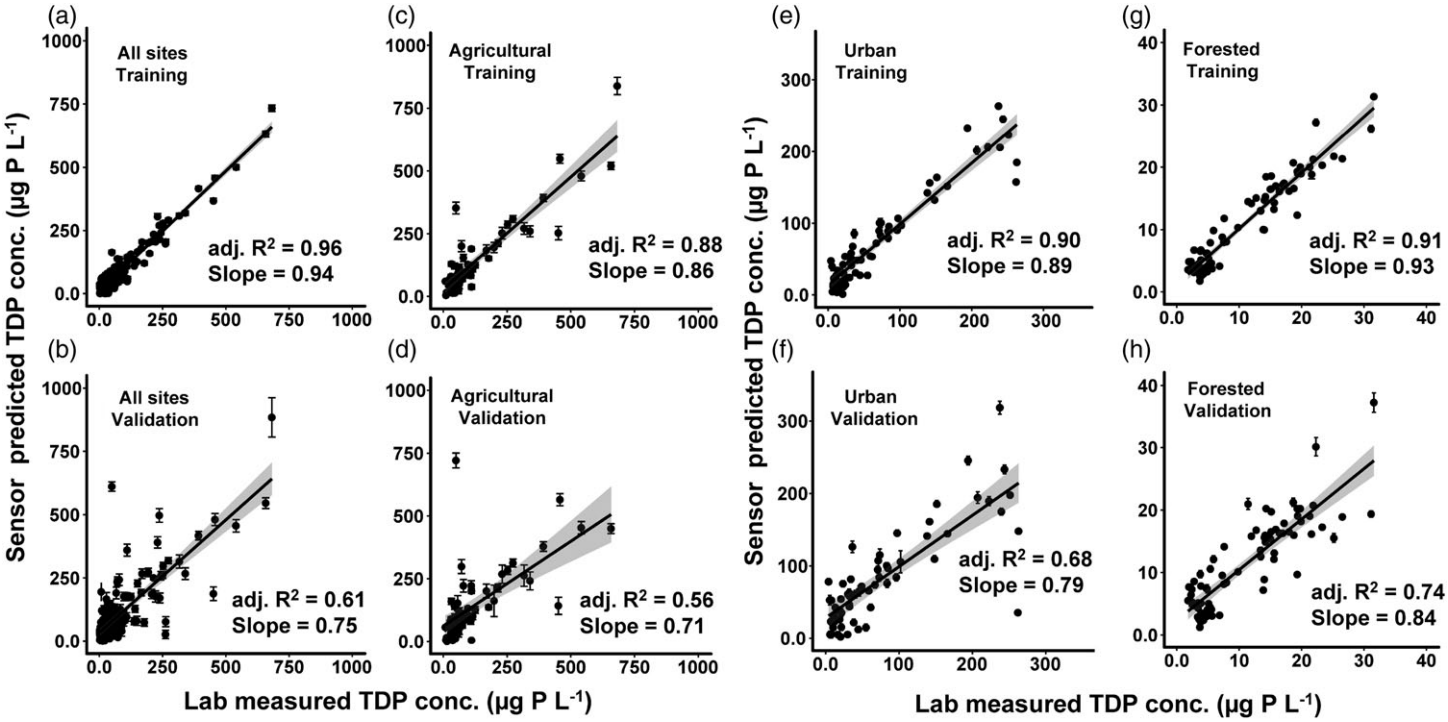

Fig. 6. Bootstrap TDP training and validation plots for (**a**, **b**) all combined, (**c**, **d**) agricultural, (**e**, **f**) urban, and (**g**, **h**) forested sites. All correlations were statistically significant ($p < 0.001$). Shading represents 90% confidence intervals. Error bars represent one standard deviation for the predictions over 1000 bootstrap iterations. Note that error bars are present for all points but may not be visible and that scales differ among plots.

correlations from the bootstrap validation method explained approximately two-thirds of the variance in SRP concentration (adj. $R^2 = 0.68$; $p < 0.001$) (Table 3; Fig. 7a,b). Separating datasets by site improved validation accuracy for the urban and forested sites, but did not improve validation accuracy at the agriculture site (Fig. 7c–h). As with TDP models, no discernable patterns could be found by plotting SRP model residuals by lab-measured value, turbidity, and discharge (Supporting Information Fig. C).

Discussion

UV-Visible spectra as proxies for phosphorus fraction concentrations

Integrated results from this study suggest that in situ UV-Visible spectrophotometers can concurrently predict the concentration and distribution of the phosphorus fractions (TP, TDP, and SRP) at a high frequency and with modest and variable accuracy that may be suitable for some applications (Fig. 8). Model goodness of fit statistics for these fractions are among the most favorable published for other proxy models. Accuracy limitations remain, however, as RMSE statistics were relatively high compared to median concentration values at our study sites. These analyses indicate that in streams draining watersheds of different primary LULCs and varying seasonal and event conditions, the measured UV-Visible absorbance spectra covaried with a suite of constituents that varied in proportion with phosphorus fractions of interest. The degree to which phosphorus fraction concentration correlates with components of the absorbance spectra can be site-specific and may vary by fraction and/or dominant biogeochemical processes and hydrologic pathways within a particular catchment. In the following discussion, we focus on the strengths and limitations of this approach, and make recommendations for how researchers and water resource managers can use this technology for monitoring phosphorus.

Models to predict TP concentrations using all data available explained a relatively high proportion of the variance, but had RMSE values that were higher than the median concentrations at the urban and forested sites (Fig. 5a,b). Site-specific models had higher accuracy but lower predictive power for the forested site where phosphorus concentrations were lower. We found that models from UV-Visible spectra explained more of the variance in TP concentration than multiple linear regression models using turbidity and logarithmically transformed discharge. The method for TP prediction demonstrated here may be best used in agricultural areas or other sites with elevated TP phosphorus concentrations; these areas may also be where this technology could be most useful for informing management goals.

The UV-Visible spectra were used to predict TDP and SRP concentrations with a greater proportion of variance explained than any other models based on a high-frequency method

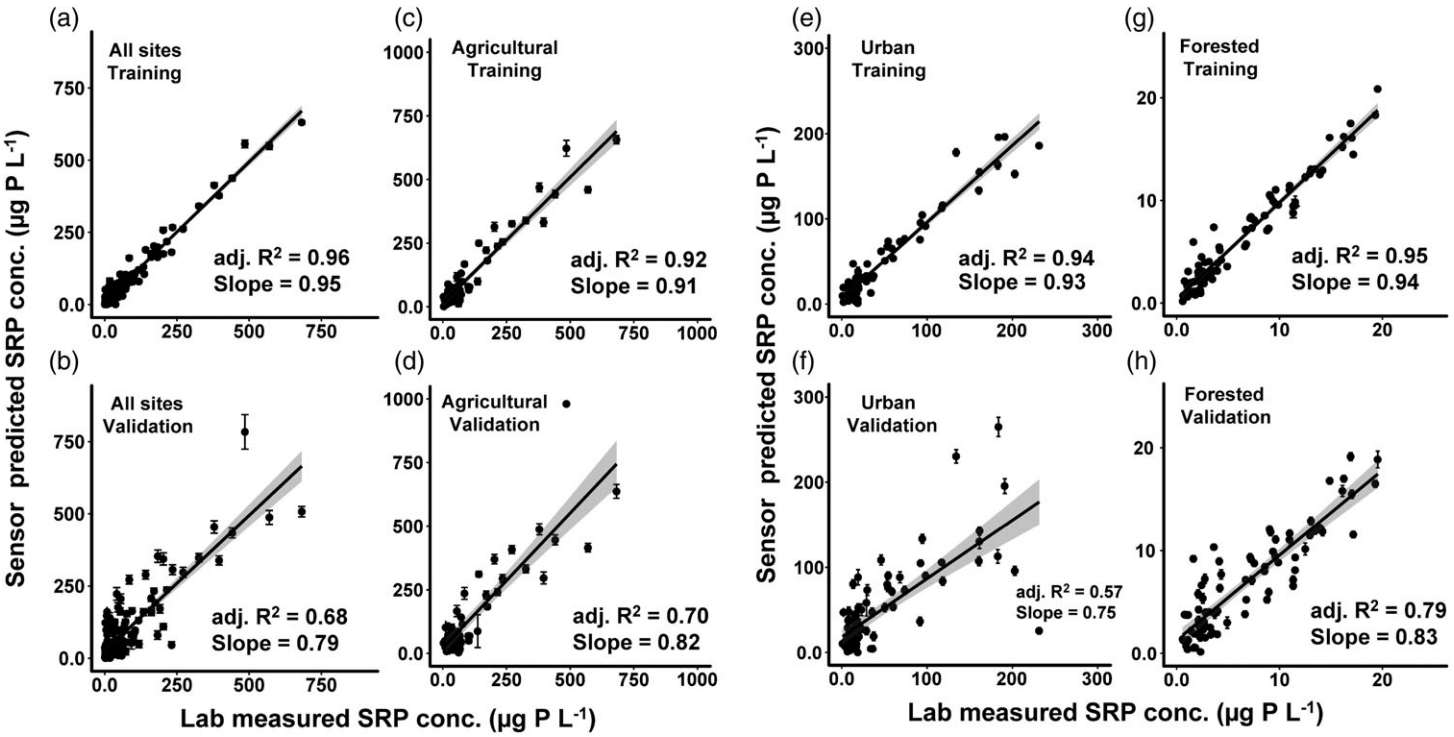

Fig. 7. Bootstrap SRP training and validation plots for (**a**, **b**) all combined, (**c**, **d**) agricultural, (**e**, **f**) urban, and (**g**, **h**) forested sites. All correlations were statistically significant ($p < 0.001$). Shading represents 90% confidence intervals. Error bars represent one standard deviation for the predictions over 1000 bootstrap iterations. Note that error bars are present for all points but may not be visible and that scales differ among plots.

known to the authors, though RMSE values indicate limited accuracy for low concentrations. The proportion of variance explained suggests that this method is a useful approach to characterize TDP and SRP concentrations, particularly during hot moments for phosphorus transport when concentrations can become elevated (e.g., Underwood et al. 2017). The high bioavailability of dissolved phosphorus fractions makes the unique ability of this approach to model both the TDP and SRP fractions particularly useful. Furthermore, the necessity of site-specific models suggests that sources and pools of dissolved phosphorus likely differ among sites, and that phosphorus fractions covary with different components of the water matrix in contrasting LULCs. In the forested site, organic and inorganic P cycling is primarily from parent material weathering and ecosystem cycling (Likens 2013). While these processes also occur in the urban and agricultural systems, fertilizer amendments and other human activities in urban and agricultural catchments add additional organic and inorganic phosphorus (Daloğlu et al. 2012). Since UV-Visible spectrophotometers have been shown to accurately model dissolved organic carbon concentration (Ruhala and Zarnetske 2017; Vaughan et al. 2017), variance in the phosphorus models may be explained by the presence of organically bound phosphorus. These pools are likely to differ among LULCs, which have very different sources and pools of organic matter (Sickman et al. 2007; Wilson and Xenopoulos 2009). These differences are often more pronounced during storm events, when rapid changes in hydrology cause changes in connectivity of differing source areas (e.g., edge of a row crop field vs. a suburban development) to streams.

Site-specific TDP and SRP concentration models performed better than models based on data from all sites for each solute. Therefore, each stream has a distinct relationship between the portion of the aqueous matrix that absorbs UV-Visible light and dissolved phosphorus fraction concentrations (Fig. 3). The PLSR method tested here relies on the shape of each UV-Visible spectrum curve to determine the phosphorus fraction concentration rather than a narrow wavelength range of absolute absorbance values. This result indicates that the method uses these distinct relationships between dissolved phosphorus fractions and various aqueous and solid constituents that absorb light across the UV-Visible range that manifest in variable absorbance spectra. While it seems that site-specific calibrations were optimal in this study, it is not yet known whether these relationship differences are due to LULC alone, or whether sites with similar LULCs could have different relationships. Further testing at several agricultural sites, for example, would help determine whether models should be strictly site-specific, or if LULC-specific models could suffice.

Comparison to other approaches

Several other studies have attempted to relate phosphorus fraction concentrations with parameters that are easier, cheaper, and faster to measure than direct measurement with

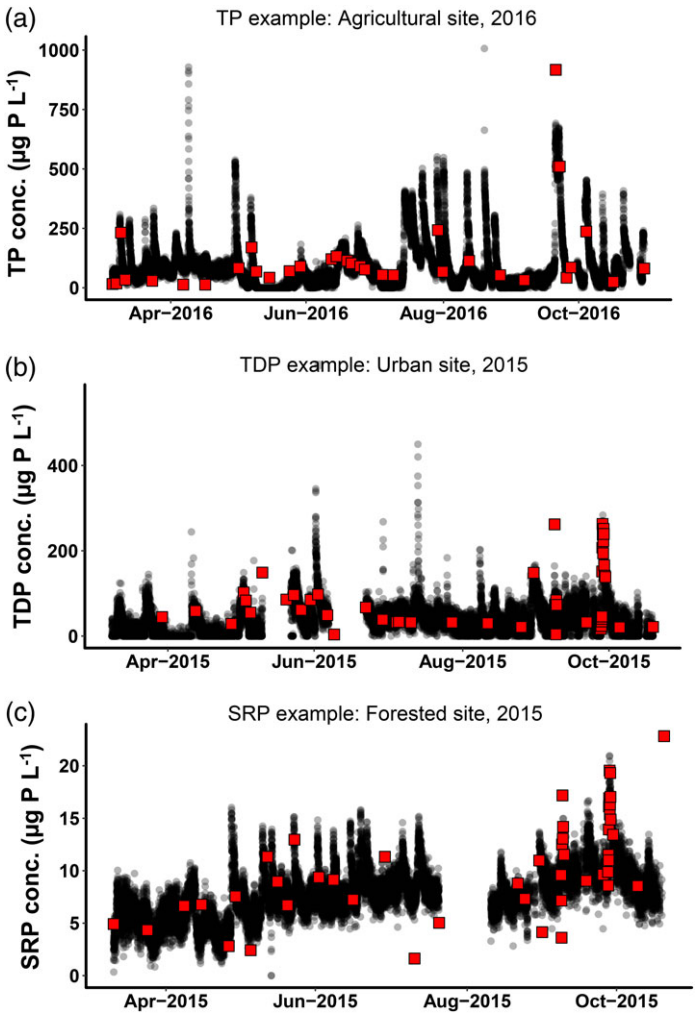

Fig. 8. Examples of modeled 15-min phosphorus fraction concentrations using UV-Visible spectra (black dots) and lab-measured values (red squares) for (**a**) TP at the agricultural site, (**b**) TDP at the urban site, and (**c**) SRP at the forested site.

wet chemistry lab techniques. Other studies showed that roughly 60–95% of the variance in TP concentration can be explained by turbidity or discharge, or these variables in combination with other proxy variables (Table 4). Results from these studies are derived from models that were based on the predicted data that were used to build the models originally. Thus, their results are most comparable to the training sets reported here, with the difference that 100% of the measurements were commonly used in these other studies, while 85% of the data was used in our training sets. The variance explained in training sets in this study was near or above 90% for all models, exceeding that of most other models reported in the literature. In addition, when we attempted to use logarithmically transformed discharge and turbidity measurements to predict TP concentrations, we found that these models explained a lower proportion of the variance compared with models based on UV-Visible absorbance spectra. Only 55% of the variance in TP concentrations could be explained by a combination of discharge and turbidity, while 96% of the variance in TP concentrations could be explained by the training model derived from the UV-Visible spectra. The higher proportion of variance explained by the spectrophotometric proxies compared to the discharge and turbidity proxies may be because our systems are smaller and more susceptible to short-timescale hysteresis-related changes in the relationship between these variables. Other studies that reported a higher proportion of variance in TP concentration explained were in rivers with larger watershed areas than we investigated.

Few studies have investigated the relationship between TDP and SRP concentrations and proxy variables. Underwood et al. (2017) recently used Bayesian linear regression to correlate TDP to discharge and identify operational thresholds where shifts in these relationships occur. More often, TP is predicted using a proxy such as turbidity or discharge, and a percentage of TDP or SRP to TP from a subset of samples is applied uniformly to estimate TDP or SRP loads (e.g., Johnes 2007). We observed that the ratios of lab-measured TDP to TP and SRP to TP varied considerably (Fig. 4), so assuming a constant relationship between these fractions would lead to considerable errors in phosphorus fraction load estimation in the systems studied here. Stubblefield et al. (2007) found no correlation between SRP concentration and turbidity measurements in a subalpine forested stream where the discharge-weighted mean SRP concentration was 8.7% of the TP concentration. Using similar methods to this study, Birgand et al. (2016) found that UV-Visible absorbance explained 89% of the variance in observed SRP concentrations in a eutrophic drinking water reservoir, which is similar to the model results for our SRP model training sets. Besides environmental setting, that study differed from this one in a few notable ways: seven components were used with 36 samples to develop a calibration (~ 20% rather than ~ 10% of the number of observations used here), SRP concentrations ranged from 3.5 μg P L^{-1} to 10 μg P L^{-1} (a narrower range than our sites), and no model validation results were reported.

For high-frequency water quality measurements, in situ UV-Visible spectrophotometers have several advantages and some limitations. Advantages include the ability to measure multiple parameters concurrently and rapidly with no reagents, and to deploy sensors for continuous monitoring of baseflow and larger episodic events. There are field-robust models available that have few moving parts to service. However, limitations include their high cost (currently greater than US$15,000), which can be prohibitive. As discussed above, UV-Visible spectrophotometer accuracy for phosphorus fraction concentrations and some other analytes may not be acceptable for some applications, particularly at relatively low-phosphorus concentrations. For the s::can sensor used here, two further limitations were on-board memory storage and power draw. On-board memory capacity allowed storage of roughly 15 d of observations at a 15-min sampling interval. It

Table 4. Summary of selected studies that related phosphorus fraction concentration to other water quality parameters.

Location	Setting	Watershed landscape/ characterization	Watershed area or waterbody size (km^2)	Proxy variable(s)	Statistical method	Observations	R^2	Accuracy	Reference
Total phosphorus									
Chesapeake Bay watershed, U.S.A.	River	Agricultural	136	ln(Discharge), water temperature	Multiple linear regression	38	0.82	Mallows' Cp = 2.71	Hyer et al. (2016)
Chesapeake Bay watershed, U.S.A.	River	Agricultural	95	ln(Discharge), dissolved oxygen, ln(Turbidity)	Multiple linear regression	32	0.96	Mallows' Cp = 2.09	Hyer et al. (2016)
North Carolina, U.S.A.	Brackish marsh	Constructed brackish marsh, agricultural	NA	UV-Visible absorbance	Partial least squares regression	NA	0.73*	RMSE = 23 μ g P L^{-1}	Etheridge et al. (2014)
Vermont, U.S.A. and Québec, CA	River	Agricultural	92	Acoustic Doppler profiler backscatter	Linear regression with log-transformed data and Duan correction	317	0.67	NA	Schuett and Bowden (2014)
Lake Tahoe Basin, California, U.S.A.	River	Subalpine forest	25	Turbidity	Linear regression	117	0.62	NA	Stubblefield et al. (2007)
Lake Tahoe Basin, California, U.S.A.	River	Subalpine forest	29.5	Turbidity	Linear regression	51	0.83	NA	Stubblefield et al. (2007)
Australia	River	Mixed landscape	5000	Turbidity	Linear regression	NA	0.90	NA	Grayson et al. (1996)
Soluble reactive phosphorus (PO_4^{3-})									
West Virginia, U.S.A.	Reservoir	Eutrophic drinking water reservoir	0.12	UV-Visible absorbance	Partial least squares regression	36	0.89	2X residual SE = 1.08 μg P L^{-1}	Birgand et al. (2016)
North Carolina, U.S.A.	Brackish marsh	Constructed brackish marsh, agricultural	NA	UV-Visible absorbance	Partial least squares regression	NA	0.66*	RMSE = 10 μg P L^{-1}	Etheridge et al. (2014)
Lake Tahoe basin, California, U.S.A.	River	Subalpine forest	25	Turbidity	Linear regression	NA	No correlation	NA	Stubblefield et al. (2007)

*Model validation was performed but not reported.

was a challenge at times to provide necessary power to the sensors when light to our solar panel array was limited by season and/or tree canopy cover.

Instruments that use wet chemistry techniques to measure SRP concentrations directly with the ascorbic acid method in situ have recently become available. For example, the Cycle-PO4 instrument (Wetlabs, Philomath, Oregon, U.S.A.) makes direct measurements of SRP concentration with onboard standard checks, which may produce a more accurate estimate of SRP concentration. Results from Cohen et al. (2013) and Sherson et al. (2015) suggest that the Cycle-PO4 measures SRP more accurately than the UV-Visible spectrophotometers tested here at low concentrations. The Systea WIZ probe (Systea, Anagni, Italy) has a similar method to the Cycle-PO4 and also tested relatively well for predicting SRP concentration in recent evaluations (Copetti et al. 2017; Johengen et al. 2017). However, these instruments have several components such as pumps, switches, and filters that are prone to malfunction; they use reagents that generate hazardous waste; and they are more prone to fouling (Pellerin et al. 2016). Both the Cycle-PO4 and the Systea-PO4 have limited capacity to measure elevated SRP concentrations, such as those found in our agricultural and urban sites. The Cycle-PO4 is specified for SRP concentrations of 0–300 μg P L^{-1}, and the Systea-PO4 was shown to have limited accuracy for concentrations above 40 μg P L^{-1}. In addition, limited reagent lifetime and sampling frequency precludes the Cycle-PO4 sensor from long-term deployments in remote or rapidly changing environments. The sampling frequency also limits its application for vertical or lateral profiling, where UV-Visible spectrophotometers can be useful. A UV-Visible spectrophotometer is preferable to an in situ wet chemistry instrument if researchers would benefit from concurrent measurements of multiple phosphorus fractions (TP, TDP, and SRP), nitrate (e.g., Rode et al. 2016), dissolved organic carbon (e.g., Ruhala and Zarnetske 2017), and other potential analytes (Birgand et al. 2016) with a single instrument. This concurrent measurement advantage may be the greatest strength of the UV-Visible spectra approach, though building a calibration dataset comes with a considerable cost that will depend on site-specific considerations.

To the authors' knowledge, this study is the first to use a rigorous bootstrap validation technique to investigate how well models predict phosphorus fraction concentrations where lab-measured values are not available. Etheridge et al. (2014) and Vaughan et al. (2017) are the only studies we are aware of that test nutrient prediction models by withholding a portion of each calibration dataset (equal to 10% in those studies). This study takes the next step in repeating these validations many times with random observation set selection to reduce sampling error when selecting the 15% to withhold. Our results reflect the expectation that validation models explain less variance than training models (Table 3) and demonstrate that method performance may have been inflated by reporting of training sets alone in previous studies. Validation sets are standard for larger-scale models in other scientific disciplines such as global climate general circulation models (Chervin 1981; Flato et al. 2013), though the exercise is valuable when using a model to predict a dependent variable at any scale. We recommend that future studies using high-frequency water quality sensors perform model validation with bootstrapping to more rigorously estimate uncertainty for new analyte concentration predictions. This approach is particularly useful when developing models that rely on absorbance spectra derived from in situ spectrophotometers to project the concentration of solutes such as dissolved phosphorus fractions that do not directly absorb light in the UV-Visible spectrum.

Implications for application in watershed monitoring

The advantage of high-frequency water quality data is generally twofold: it can reveal short-timescale effects previously invisible to researchers, and it can aid in more accurate load estimation. Because our results indicate that UV-Visible absorbance is generally sensitive to changes in phosphorus fraction concentrations (models had acceptable coefficients of determination), but had relatively low accuracy (models had relatively high RMSE values), we suggest that in situ spectrophotometers are best applied to understanding short-timescale phosphorus dynamics, especially in systems with relatively high-phosphorus fraction concentrations. Depending on site-specific model performance, this technology may be suited to provide valuable, yet possibly semi-quantitative information about phosphorus fraction dynamics during storms or diel cycles, illuminating potential nutrient sources and biological processes. This technique could be especially informative when developed in combination with models for other useful parameters (e.g., nitrate and dissolved organic carbon).

The relatively high RMSE value to median concentrations ratios found here suggest that phosphorus load estimates calculated with this method may have substantial uncertainty, unless site-specific models elsewhere show improved accuracy. Optimal models to predict TP concentration had RMSE values that were 75–80% of median TP concentrations. This ratio is an indication of the level of uncertainty a load estimate may have, though actual uncertainty would depend on annual hydrologic conditions and site-specific factors.

Conclusions and recommendations

We have shown that UV-Visible spectra collected by in situ spectrophotometric sensors can be used to simultaneously predict TP, TDP, and SRP concentrations in many situations. For our sites, the ratios of TDP to TP and SRP to TP varied notably, so that if high-frequency measurements of TDP and SRP were of main interest in a study or management decision, the use of in situ spectrophotometers is clearly warranted. Since these sensors also measure turbidity, nitrate, and dissolved organic carbon concentrations, there is the capability to measure diverse chemical constituents concurrently. If estimates

for these other parameters are a primary monitoring goal, phosphorus fractions model development could be a relatively low-risk, low-cost addition.

This technology is best suited to sites with elevated TP concentrations if TP concentrations are the primary fraction of interest. We recommend that all models be checked to determine if separating data by site improves or weakens model performance. When using the PLSR method, we recommend following Mevik et al. (2015) to use the number of components equal to ~ 10% of observations, as a higher percentage of components can lead to over-parameterization. Over-parameterization may lead to more favorable training model statistics, but also to weaker validation model performance statistics, and noisier and less accurate time series prediction. The success of this method may be influenced by the number and variety of grab samples that can be attained, analyzed, and incorporated into prediction models. We recommend that users of this technology take care to obtain grab samples as close in time and space to the sensor measurement as possible to obtain a reliable calibration.

There has been significant effort to create "global calibrations" or calibration "libraries" for various predictive proxies and predicted constituents (e.g., Shepherd and Walsh 2002). Although this type of effort is beyond the scope of this study, our results indicate that that common models for phosphorus fraction concentrations were not preferable to site-specific models for three sites with variable LULC. Future work is necessary to rule out the possibility of a more extensive library to explain a greater amount of variance across multiple types of sites and water matrices.

The number of samples needed to develop useful models to predict phosphorus fraction concentrations using UV-Visible spectra will be dependent on many factors that will likely be site-specific. For example, the greater the variability in the concentration at the monitoring location, the more samples will be needed to form an adequate predictive model. Although evaluating these criteria will depend on subjective expert opinion, researcher geochemical/hydrologic intuition, and available observational data prior to sensor deployment, we suggest that an adequate PLSR model must meet the following conditions:

1. The PLSR model has a validated accuracy and goodness of fit that is acceptable for the application.
2. The number of observations is equal to or greater than 10 times the number of components in the PLSR model.
3. The range of the sampled concentrations is approximately equal to the range of concentrations likely to occur at the site.
4. Samples were collected during times representative of the various conditions at the site (e.g., baseflow, rising and falling limbs of storms, seasonal conditions, nutrient amendment schedules, biological hot moments, *see* Fig. 2).

As use of in situ optical spectrophotometers increases, researchers and managers will gain a better picture of their performance to measure several water quality parameters. In the foreseeable future, this type of instrumentation may extend our ability to monitor critical nutrients at times and places that would be difficult to sample in any other way. Results presented in this work also indicate that with further study in a more diverse set of environments, phosphorus fractions may be monitored with increasing reliability to inform watershed management goals.

References

Aber, J. D. 1997. Why don't we believe the models? Bull. Ecol. Soc. Am. **78**: 232–233.

Avagyan, A., B. R. K. Runkle, and L. Kutzbach. 2014. Application of high-resolution spectral absorbance measurements to determine dissolved organic carbon concentration in remote areas. J. Hydrol. **517**: 435–446. doi:10.1016/j.jhydrol.2014.05.060

Bieroza, M. Z., and A. L. Heathwaite. 2015. Seasonal variation in phosphorus concentration–discharge hysteresis inferred from high-frequency in situ monitoring. J. Hydrol. **524**: 333–347. doi:10.1016/j.jhydrol.2015.02.036

Birgand, F., K. Aveni-Deforge, B. Smith, B. Maxwell, M. Horstman, A. B. Gerling, and C. C. Carey. 2016. First report of a novel multiplexer pumping system coupled to a water quality probe to collect high temporal frequency in situ water chemistry measurements at multiple sites. Limnol. Oceanogr.: Methods **14**: 767–783. doi:10.1002/lom3.10122

Carey, R. O., W. M. Wollheim, G. K. Mulukutla, and M. M. Mineau. 2014. Characterizing storm-event nitrate fluxes in a fifth order suburbanizing watershed using in situ sensors. Environ. Sci. Technol. **48**: 7756–7765. doi:10.1021/es500252j

Carpenter, S. R., N. F. Caraco, D. L. Correll, R. W. Howarth, A. N. Sharpley, and V. H. Smith. 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecol. Appl. **8**: 559–568. doi:10.1890/1051-0761(1998)008[0559:NPOSWW]2.0.CO;2

Chervin, R. M. 1981. On the comparison of observed and GCM simulated climate ensembles. J. Atmos. Sci. **38**: 885–901. doi:10.1175/1520-0469(1981)038<0885:OTCOOA>2.0.CO;2

Cohen, M. J., M. J. Kurz, J. B. Heffernan, J. B. Martin, R. L. Douglass, C. R. Foster, and R. G. Thomas. 2013. Diel phosphorus variation and the stoichiometry of ecosystem metabolism in a large spring-fed river. Ecol. Monogr. **83**: 155–176. doi:10.1890/12-1497.1

Conley, D. J., H. W. Paerl, R. W. Howarth, D. F. Boesch, S. P. Seitzinger, K. E. Havens, C. Lancelot, and G. E. Likens. 2009. Controlling eutrophication: Nitrogen and phosphorus. Science **323**: 1014–1015. doi:10.1126/science.1167755

Copetti, D., L. Valsecchi, A. G. Capodaglio, and G. Tartari. 2017. Direct measurement of nutrient concentrations in freshwaters with a miniaturized analytical probe: Evaluation

and validation. Environ. Monit. Assess. **189**: 144. doi:10.1007/s10661-017-5847-0

Correll, D. L. 1998. The role of phosphorus in the eutrophication of receiving waters: A review. J. Environ. Qual. **27**: 261–266. doi:10.2134/jeq1998.00472425002700020004x

Correll, D. L., T. E. Jordan, and D. E. Weller. 1999. Transport of nitrogen and phosphorus from rhode river watersheds during storm events. Water Resour. Res. **35**: 2513–2521. doi:10.1029/1999WR900058

Daloğlu, I., K. H. Cho, and D. Scavia. 2012. Evaluating causes of trends in long-term dissolved reactive phosphorus loads to Lake Erie. Environ. Sci. Technol. **46**: 10660–10666. doi:10.1021/es302315d

Dhillon, G. S., and S. Inamdar. 2013. Extreme storms and changes in particulate and dissolved organic carbon in runoff: Entering uncharted waters? Geophys. Res. Lett. **40**: 1322–1327. doi:10.1002/grl.50306

Djodjic, F., H. Montas, A. Shirmohammadi, L. Bergström, and B. Ulén. 2002. A decision support system for phosphorus management at a watershed scale. J. Environ. Qual. **31**: 937–945. doi:10.2134/jeq2002.9370

Dodd, R. J., and A. N. Sharpley. 2016. Conservation practice effectiveness and adoption: Unintended consequences and implications for sustainable phosphorus management. Nutr. Cycl. Agroecosyst. **104**: 373–392. doi:10.1007/s10705-015-9748-8

Etheridge, J. R., F. Birgand, J. A. Osborne, C. L. Osburn, M. R. Burchell II, and J. Irving. 2014. Using in situ ultraviolet-visual spectroscopy to measure nitrogen, carbon, phosphorus, and suspended solids concentrations at a high frequency in a brackish tidal marsh. Limnol. Oceanogr.: Methods **12**: 10–22. doi:10.4319/lom.2014.12.10

Fichot, C. G., and R. Benner. 2011. A novel method to estimate DOC concentrations from CDOM absorption coefficients in coastal waters. Geophys. Res. Lett. **38**. doi:10.1029/2010GL046152

Flato, G., and others. 2013. Evaluation of climate models, p. 741–866. *In* T. F. Stocker and others. [eds.] Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the Intergovernmental Panel on Climate Change. Cambridge Univ. Press.

Giles, C. D., L. G. Lee, B. J. Cade-Menun, J. E. Hill, P. D. F. Isles, A. W. Schroth, and G. K. Druschel. 2015. Characterization of organic phosphorus form and bioavailability in lake sediments using ^{31}P nuclear magnetic resonance and enzymatic hydrolysis. J. Environ. Qual. **44**: 882–894. doi:10.2134/jeq2014.06.0273

Grayson, R. B., B. L. Finlayson, C. J. Gippel, and B. T. Hart. 1996. The potential of field turbidity measurements for the computation of total phosphorus and suspended solids loads. J. Environ. Manage. **47**: 257–267. doi:10.1006/jema.1996.0051

Guo, Y., M. Markus, and M. Demissie. 2002. Uncertainty of nitrate-N load computations for agricultural watersheds. Water Resour. Res. **38**: 3-1–3-12. doi:10.1029/2001WR001149

Heffernan, J. B., and M. J. Cohen. 2010. Direct and indirect coupling of primary production and diel nitrate dynamics in a subtropical spring-fed river. Limnol. Oceanogr. **55**: 677–688. doi:10.4319/lo.2009.55.2.0677

Hirsch, R. M., D. L. Moyer, and S. A. Archfield. 2010. Weighted regressions on time, discharge, and season (WRTDS), with an application to Chesapeake Bay river inputs. JAWRA J. Am. Water Resour. Assoc. **46**: 857–880. doi:10.1111/j.1752-1688.2010.00482.x

Hyer, K. E., Denver, J. M., Langland, M. J., Webber, J. S., Böhlke, J. K., Hively, W. D., and Clune, J. W. 2016. Spatial and temporal variation of stream chemistry associated with contrasting geology and land-use patterns in the Chesapeake Bay watershed—summary of results from Smith Creek, Virginia; Upper Chester River, Maryland; Conewago Creek, Pennsylvania; and Difficult Run, Virginia, 2010–2013. doi:10.3133/sir20165093

Isles, P. D. F., Y. Xu, J. D. Stockwell, and A. W. Schroth. 2017. Climate-driven changes in energy and mass inputs systematically alter nutrient concentration and stoichiometry in deep and shallow regions of Lake Champlain. Biogeochemistry **133**: 201–217. doi:10.1007/s10533-017-0327-8

Jarvie, H. P., L. T. Johnson, A. N. Sharpley, D. R. Smith, D. B. Baker, T. W. Bruulsema, and R. Confesor. 2017. Increased soluble phosphorus loads to Lake Erie: Unintended consequences of conservation practices? J. Environ. Qual. **46**: 123–132. doi:10.2134/jeq2016.07.0248

Johengen, T., and others. 2017. Performance verification statment for Systea WIZ Probe phosphoate analyzer. Alliance for Coastal Technologies.

Johnes, P. J. 2007. Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density. J. Hydrol. **332**: 241–258. doi:10.1016/j.jhydrol.2006.07.006

Jordan, P., A. Arnscheidt, H. Mcgrogan, and S. Mccormick. 2007. Characterising phosphorus transfers in rural catchments using a continuous bank-side analyser. Hydrol. Earth Syst. Sci. Discuss. **11**: 372–381. doi:10.5194/hess-11-372-2007

Joung, D., and others. 2017. Winter weather and lake-watershed physical configuration drive phosphorus, iron, and manganese dynamics in water and sediment of ice-covered lakes. Limnol. Oceanogr. **62**: 1620–1635. doi:10.1002/lno.10521

Kane, D. D., J. D. Conroy, R. Peter Richards, D. B. Baker, and D. A. Culver. 2014. Re-eutrophication of Lake Erie: Correlations between tributary nutrient loads and phytoplankton biomass. J. Great Lakes Res. **40**: 496–501. doi:10.1016/j.jglr.2014.04.004

Kruskal, W. H., and W. A. Wallis. 1952. Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. **47**: 583–621. doi:10.2307/2280779

Langergraber, G., N. Fleischmann, and F. Hofstadter. 2003. A multivariate calibration procedure for UV/VIS spectrometric quantification of organic matter and nitrate in wastewater. Water Sci. Technol. **47**: 63–71. doi:10.2166/wst.2003.0086

Likens, G. E. 2013. Biogeochemistry of a forested ecosystem, p. 208. Springer.

McCarty, G. W., J. B. Reeves, V. B. Reeves, R. F. Follett, and J. M. Kimble. 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurement. Soil Sci. Soc. Am. J. **66**: 640–646. doi:10.2136/sssaj2002.6400

Medalie, L. 2016. Concentration, flux, and trend estimates with uncertainty for nutrients, chloride, and total suspended solids in tributaries of Lake Champlain, 1990–2014. doi:10.3133/ofr20161200

Mevik, B., Wehrens, R., and Liland, K. H., 2015. pls: Partial least squares and principal component regression. R package version 2.6-0. https://CRAN.R-project.org/package=pls

Musolff, A., J. H. Fleckenstein, P. S. C. Rao, and J. W. Jawitz. 2017. Emergent archetype patterns of coupled hydrologic and biogeochemical responses in catchments. Geophys. Res. Lett. **44**: 4143–4151. doi:10.1002/2017GL072630

Ohtani, K. 2000. Bootstrapping R^2 and adjusted R^2 in regression analysis. Econ. Model. **17**: 473–483. doi:10.1016/S0264-9993(99)00034-6

Parsons, T. R., Y. Maita, and C. M. Lalli. 1984. Determination of phosphate, p. 22–25. A manual of chemical and biological methods for seawater analysis. Pergamon. doi:10.1016/C2009-0-07774-5, ISBN: 978-0-08-030287-4

Pellerin, B. A., J. F. Saraceno, J. B. Shanley, S. D. Sebestyen, G. R. Aiken, W. M. Wollheim, and B. A. Bergamaschi. 2012. Taking the pulse of snowmelt: In situ sensors reveal seasonal, event and diurnal patterns of nitrate and dissolved organic matter variability in an upland forest stream. Biogeochemistry **108**: 183–198. doi:10.1007/s10533-011-9589-8

Pellerin, B. A., B. A. Bergamaschi, R. J. Gilliom, C. G. Crawford, J. Saraceno, C. P. Frederick, B. D. Downing, and J. C. Murphy. 2014. Mississippi River nitrate loads from high frequency sensor measurements and regression-based load estimation. Environ. Sci. Technol. **48**: 12612–12619. doi:10.1021/es504029c

Pellerin, B. A., B. A. Stauffer, D. A. Young, D. J. Sullivan, S. B. Bricker, M. R. Walbridge, G. A. Clyde, and D. M. Shaw. 2016. Emerging tools for continuous nutrient monitoring networks: Sensors advancing science and water resources protection. JAWRA J. Am. Water Resour. Assoc. **52**: 993–1008. doi:10.1111/1752-1688.12386

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Rieger, L., G. Langergraber, and H. Siegrist. 2006. Uncertainties of spectral in situ measurements in wastewater using different calibration approaches. Water Sci. Technol. **53**: 187–197. doi:10.2166/wst.2006.421

Rode, M., and others. 2016. Sensors in the stream: The high-frequency wave of the present. Environ. Sci. Technol. **50**: 10297–10307. doi:10.1021/acs.est.6b02155

Rosenberg, B. D., and A. W. Schroth. 2017. Coupling of reactive riverine phosphorus and iron species during hot transport moments: Impacts of land cover and seasonality. Biogeochemistry **132**: 103–122. doi:10.1007/s10533-016-0290-9

Ruhala, S. S., and J. P. Zarnetske. 2017. Using in-situ optical sensors to study dissolved organic carbon dynamics of streams and watersheds: A review. Sci. Total Environ. **575**: 713–723. doi:10.1016/j.scitotenv.2016.09.113

Sakamoto, C. M., K. S. Johnson, and L. J. Coletti. 2009. Improved algorithm for the computation of nitrate concentrations in seawater using an in situ ultraviolet spectrophotometer. Limnol. Oceanogr.: Methods **7**: 132–143. doi:10.4319/lom.2009.7.132

Saraceno, J. F., B. A. Pellerin, B. D. Downing, E. Boss, P. A. M. Bachand, and B. A. Bergamaschi. 2009. High-frequency in situ optical measurements during a storm event: Assessing relationships between dissolved organic matter, sediment concentrations, and hydrologic processes. J. Geophys. Res. Biogeosci. **114**. doi:10.1029/2009JG000989

Schroth, A. W., C. D. Giles, P. D. F. Isles, Y. Xu, Z. Perzan, and G. K. Druschel. 2015. Dynamic coupling of iron, manganese, and phosphorus behavior in water and sediment of shallow ice-covered eutrophic lakes. Environ. Sci. Technol. **49**: 9758–9767. doi:10.1021/acs.est.5b02057

Schuett, E., and Bowden, W. B. 2014. Use of acoustic Doppler current profiler data to estimate sediment and total phosphorus loads to Lake Champlain from the Rock River. Final report to the Vermont Agency of Natural Resources.

Sharpley, A. N., S. C. Chapra, R. Wedepohl, J. T. Sims, T. C. Daniel, and K. R. Reddy. 1994. Managing agricultural phosphorus for protection of surface waters: Issues and options. J. Environ. Qual. **23**: 437–451. doi:10.2134/jeq1994.00472425002300030006x

Sharpley, A. N., P. J. A. Kleinman, A. L. Heathwaite, W. J. Gburek, G. J. Folmar, and J. P. Schmidt. 2008. Phosphorus loss from an agricultural watershed as a function of storm size. J. Environ. Qual. **37**: 362–368. doi:10.2134/jeq2007.0366

Shepherd, K. D., and M. G. Walsh. 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Sci. Soc. Am. J. **66**: 988–998. doi:10.2136/sssaj2002.9880

Sherson, L. R., D. J. Van Horn, J. D. Gomez-Velez, L. J. Crossey, and C. N. Dahm. 2015. Nutrient dynamics in an alpine headwater stream: Use of continuous water quality sensors to examine responses to wildfire and precipitation events. Hydrol. Process. **29**: 3193–3207. doi:10.1002/hyp.10426

Sickman, J. O., M. J. Zanoli, and H. L. Mann. 2007. Effects of urbanization on organic carbon loads in the Sacramento River, California. Water Resour. Res. **43**: W11422. doi:10.1029/2007WR005954

Stubblefield, A. P., J. E. Reuter, R. A. Dahlgren, and C. R. Goldman. 2007. Use of turbidometry to characterize suspended sediment and phosphorus fluxes in the Lake Tahoe basin, California, USA. Hydrol. Process. **21**: 281–291. doi:10.1002/hyp.6234

Stumpf, R. P., T. T. Wynne, D. B. Baker, and G. L. Fahnenstiel. 2012. Interannual variability of cyanobacterial blooms in Lake Erie. PLoS One **7**: e42444. doi:10.1371/journal.pone.0042444

Stutter, M., J. J. C. Dawson, M. Glendell, F. Napier, J. M. Potts, J. Sample, A. Vinten, and H. Watson. 2017. Evaluating the use of in-situ turbidity measurements to quantify fluvial sediment and phosphorus concentrations and fluxes in agricultural streams. Sci. Total Environ. **607–608**: 391–402. doi:10.1016/j.scitotenv.2017.07.013

Underwood, K. L., D. M. Rizzo, A. W. Schroth, and M. M. Dewoolkar. 2017. Evaluating spatial variability in sediment and phosphorus concentration-discharge relationships using Bayesian inference and self-organizing maps. Water Resour. Res. **53**: 10293–10316. doi:10.1002/2017WR021353

Vaughan, M. C. H., and others. 2017. High-frequency dissolved organic carbon and nitrate measurements reveal differences in storm hysteresis and loading in relation to land cover and seasonality. Water Resour. Res. **53**: 5345–5363. doi:10.1002/2017WR020491

Viscarra Rossel, R. A., D. J. J. Walvoort, A. B. McBratney, L. J. Janik, and J. O. Skjemstad. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma **131**: 59–75. doi:10.1016/j.geoderma.2005.03.007

Wilson, H. F., and M. A. Xenopoulos. 2009. Effects of agricultural land use on the composition of fluvial dissolved organic matter. Nat. Geosci. **2**: 37–41. doi:10.1038/ngeo391

Acknowledgments

We thank Ryan Sleeper, Saul Blocher, Joshua Benes, JohnFranco Saraceno, François Birgand, and two anonymous reviewers for their helpful contributions to this work. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, Vermont EPSCoR, or any other supporting organization. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. This material is based upon work supported by the National Science Foundation under VTEPSCoR Grant EPS-1101317, EPS-IIA1330446, and OIA 1556770, the Vermont Water Resources and Lakes Studies Center (project 2016VT80B) which is part of the National Institutes for Water Resources, and NSF EAR Grant 1561014 to AWS.

Conflict of Interest

None declared.

Submitted 3 February 2018
Revised 17 September 2018
Accepted 19 September 2018

Associate editor: Clare Reimers