8-2016

Trend Analysis of Water Quality in Northwest Arkansas Streams Reflects Changes in the Watershed

Zachary Paul Simpson
University of Arkansas, Fayetteville

Trend Analysis of Water Quality in Northwest Arkansas Streams Reflects Changes in the Watershed

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Biological Engineering

by

Zachary Simpson
University of Arkansas
Bachelor of Science in Biological Engineering, 2014

August 2016
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

Dr. Brian E. Haggard
Thesis Director

Dr. Benjamin R. K. Runkle
Committee Member

Dr. J. Thad Scott
Committee Member

Abstract

Watershed export of nutrients, sediments, and chemicals impacts receiving waters. Changes within the watershed (e.g., anthropogenic or climatic) can alter the transport of constituents in streams. Stream monitoring is crucial for understanding these effects. This study developed a potential improvement to flow-adjusting constituent concentrations in streams, an important step of analyzing monitoring data in lotic systems for trends. The method incorporates a K-fold cross-validation procedure to optimize a model explaining the relationship between the concentration and streamflow, thus providing a valuable tool to researchers in water quality. Additionally, two case studies were conducted on watersheds located in northwest Arkansas using monitoring data collected from 2009 to 2015. The first case study focused on phosphorus concentrations in the Illinois River watershed and illustrated significant decreases in soluble reactive phosphorus following reductions of effluent phosphorus from upstream wastewater treatment plants. However, no significant trends were found in total phosphorus at the most downstream site on the Illinois River, suggesting that there are legacy sources of phosphorus remaining in the watershed. The second case study focused on nitrogen and phosphorus in the three main inflows to Beaver Lake, where primary productivity will likely cause the lake to violate its water quality standard for chlorophyll-a concentration. Data collected at two sites in Beaver Lake showed elevated chlorophyll-a concentrations and one site near Lowell, Arkansas, the location of a major drinking water supply intake, showed increasing trends from 2001 to 2015 for total nitrogen as well as chlorophyll-a. Monitoring data of the inflows illustrated the variability in hydrological and climatic factors (e.g., drought), which affects nutrient delivery to Beaver Lake. Long-term monitoring of streams in both watersheds will be crucial for understanding the processes that affect water quality and will better inform watershed management.

Acknowledgements

I thank my family for their unending support and encouragement; my friends for the good times and for keeping me humble; the biological and agricultural engineering department, including students, faculty, and staff, for their support of all graduate students and for fostering a great department; office mates, past and present; and all the folks at Arkansas Water Resources Center. Additionally, I want to thank Mansoor Leh, with whom I had several insightful discussions that really enabled me to begin to understand statistics (and R). The first chapter of this thesis developed after Mansoor introduced me to some key concepts and a nifty R package.

I especially thank my committee and adviser: Ben Runkle, Thad Scott, and Brian Haggard. They are all wonderful teachers, as well as mentors, who have developed me as a student and researcher. I enjoyed every meeting and class.

Last, I want to salute the two 'fuels' that propelled me through much of this work: black coffee and black metal.

Dedication

To my father, Bernard Simpson Jr, and grandfather, Bernard Simpson Sr.

Table of Contents

Table of Figures

List of Publishable Papers

- Simpson, Z.P. and B.E. Haggard. 2016. An optimized procedure for flow-adjustment of constituent concentrations for trend analysis. (Intended Journal) Journal of Hydrology. In preparation. (Thesis Chapter I)
- Simpson, Z.P. and B.E. Haggard. 2016. Are phosphorus concentrations still declining in the Illinois River Watershed, northwest Arkansas? (Intended Journal) Agricultural and Environmental Letters. In preparation. (Thesis Chapter II)
- Simpson, Z.P., J.T. Scott, and B.E. Haggard. 2016. Relating water quality in Beaver Lake to nutrient trends in watershed inflows. (Intended Journal) Journal of Environmental Quality. In preparation. (Thesis Chapter III)

Introduction

Stream water quality is a valuable resource and multiple programs, agencies, and others seek to protect or improve water quality. For example, section 319 of the Clean Water Act established state programs to manage nonpoint source (NPS) pollution in navigable waters. In order to assess any impact that these programs have on water quality, monitoring programs are required. Monitoring can not only inform us of what the current conditions are, but give an idea of the dynamics at play within the watershed (e.g., seasonality, changes following an event in the watershed, etc.). One of the tools that can be used with stream monitoring data is trend analysis (Hirsch et al., 1991).

Trend analysis finds changes in concentrations of a constituent of interest, where the manifestation of this change over time can be monotonic (i.e., only increasing or only decreasing with time). To detect monotonic trends, exogenous variables in the data will need to be accounted for (e.g., seasonality and streamflow). Often, streamflow (Q) contributes the largest portion of variation in concentrations (Helsel and Hirsch, 2002). For example, phosphorus concentrations in the Illinois River in northwest Arkansas decreased with increased base flow, illustrating dilution of point sources, but then increased with increased runoff during storm events (Green and Haggard, 2001).

One method for modelling this relationship between the concentration and Q (termed as flow-adjustment) employs locally weighted regression (LOESS; Cleveland, 1979). The residuals from a LOESS fit to the data, referred to as flow-adjusted concentrations (FACs), can then be plotted against time to detect trends (Helsel and Hirsch, 2002; Bekele and McFarland, 2004). However, studies using this technique have often used default settings on LOESS, primarily its smoothing parameter (f), in order to flow-adjust concentrations (e.g., Bekele and McFarland, 2004; White et al., 2004). The first chapter of this thesis focused on developing a method to statistically optimize *f*, so that LOESS has the smoothest possible fit to the data while still capturing the important characteristics of the relationship (Jacoby, 2000).

In addition to developing a potential improvement to trend analysis methods, this thesis conducted two case studies using monitoring data from northwest Arkansas where two priority watersheds are located: the Illinois River Watershed (IRW) and the Upper White River Basin (UWRB). Both the IRW and UWRB have been targeted for nutrient management, where nitrogen (N) and phosphorus (P) concentrations are a primary concern. Trend analysis was employed in both studies to relate events in the watershed (nutrient management, changes in climate, etc.) to possible changes in stream water quality.

The IRW, a trans-boundary watershed in Arkansas and Oklahoma, has been the focus of nutrient management due to lawsuits, watershed planning, total maximum daily loads (TMDLs) implementation, and other activities. Primarily, P is the constituent of concern, and multiple sources have been targeted. A major NPS source of P is land-applied poultry litter, which has seen improvements in management over the past decade (e.g., Sharpley et al., 2003). Additionally, wastewater treatment plants (WWTPs) in the IRW have made significant reductions in effluent P concentrations, which has resulted in decreases of stream P (Haggard, 2010; Scott et al., 2011). The second chapter of this thesis examined P concentrations in the IRW from 2009 to 2015 in order to determine whether P concentrations have continued to decrease.

The UWRB contains Beaver Lake, a primary drinking water source for northwest Arkansas. Recently, a standard for chlorophyll-a has been implemented at Beaver Lake in order to protect the lake's uses from excess algal growth (Scott and Haggard, 2015). However, it is likely that

Beaver Lake will violate this standard (Scott and Haggard, 2015), thus making UWRB a potential focus for nutrient loadings. The third chapter of this thesis examined the inflows to Beaver Lake from 2009 to 2015 for both N and P using trend analysis methods. Additionally, lake data (chlorophyll-a, N and P) from two sites on Beaver Lake were analyzed to relate changes in stream water quality to conditions in the lake.

Works Cited

Bekele, A. and A. McFarland. 2004. Regression-based flow adjustment procedures for trend analysis of water quality data. Transactions of the American Society of Agricultural Engineers 47(4):1093-1104.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 74(368): 829-836.

Green, W.R. and B.E. Haggard. 2001. Phosphorus and nitrogen concentrations and loads at Illinois River south of Siloam Springs, Arkansas, 1997-1999. United States Geological Survey. Water-Resources Investigations Report 01-4217.

Haggard, B.E. 2010. Phosphorus concentrations, loads, and sources within the Illinois River drainage area, northwest Arkansas, 1997-2008. Journal of Environmental Quality 39: 2113-2120.

Helsel, D.R., and R.M. Hirsch. 2002. Statistical methods in water resources. In *Hydrologic Analysis and Interpretation,* (pp. 1-510). United States Geological Survey.

Hirsch, R.M., R.B. Alexander, and R.A. Smith. 1991. Selection of methods for the detection and estimation of trends in water quality. Water Resources Research 27(5): 803-813.

Scott, J.T., B.E. Haggard, A.N. Sharpley, and J.J. Romeis. 2011. Change point analysis of phosphorus trends in the Illinois River (Oklahoma) demonstrates the effects of watershed management. Journal of Environmental Quality 40: 1249 - 1256.

Scott, J.T. and B.E. Haggard. 2015. Implementing effects-based water quality criteria for eutrophication in Beaver Lake, Arkansas: Linking standard development and assessment methodology. Journal of Environmental Quality 44(5): 1503-1512.

Sharpley, A.N., J.L. Weld, D.B. Beegle, P.J.A. Kleinman, W.J. Gburek, P.A. Moore Jr., and G. Mullins. 2003. Development of phosphorus indices for nutrient management planning strategies in the United States. Journal of Soil and Water Conservation 58(3): 137 - 152.

White, K.L., B.E. Haggard, and I. Chaubey. 2004. Water quality at the Buffalo National River, Arkansas, 1991-2001. Transactions of the American Society of Agricultural Engineers 47(2): 407-417.

I. An Optimized Procedure for Flow-adjustment of Constituent Concentrations for Trend Analysis

Abstract

Trend analysis of stream constituent concentrations requires adjustment for exogenous variables like discharge, because concentrations often have variable relations with flow. Trend analyses in stream water quality studies normally require an accurate characterization of the relationship between the constituent and streamflow. One popular method, locally weighted regression (LOESS), provides an effective means for flow-adjusting concentrations. However, the smoothing parameter (f), which exerts the most control over the LOESS fit, needs to be selected for each dataset analyzed to avoid over-fitting or over-smoothing the data. This study provides a robust method for determining the optimal f value (f_{opt}) for each dataset via a K-fold cross-validation procedure that minimizes prediction error in LOESS. The method is developed by analyzing datasets of 7 different constituents across 17 sites from a stream monitoring program in northwest Arkansas (USA). We recommend using 10 iterations of 10 fold cross-validation (10 x 10 CV) in order to select f_{opt} when flow-adjusting water quality data with LOESS.

Introduction

Nonpoint source (NPS) pollution is one of the greatest challenges to water quality. To mitigate the problem, section 319 of the Clean Water Act has established state programs to manage NPS pollution in navigable waters. A critical part of assessing the effect of these programs is monitoring, where monitoring programs seek to answer this question, ‘Is water quality getting better or worse?’ A decreasing pollutant concentration within a stream could reflect the implementation of best management practices on the landscape (e.g., Miltner, 2015) or a reduction in point sources (e.g., Haggard, 2010; Scott et al., 2011). On the contrary, an increasing concentration may reflect poor land management in the watershed (Carpenter et al., 1998) or increased erosion risk due to change of land use (Leh et al., 2011). The manifestation of this change in concentrations can be monotonic, where the change over time is continuous and does not reverse (Hirsch et al., 1991).

To detect these monotonic trends, similar to other water quality data analyses like load estimation, multiple years of data are needed (Harmel et al., 2006). This time period introduces variation in concentrations due to exogenous variables (e.g., hydrologic and climatic variability) which have to be accounted for in order to clearly see how concentrations are changing with time. Often, streamflow (Q) contributes the largest portion of variation in concentrations, especially for nutrients and suspended sediment (Helsel and Hirsch, 2002; Lettenmaier et al., 1991). Streamflow can affect concentrations through dilution (i.e., concentration is reduced with increasing Q) or through runoff (i.e., overland flow delivers more of the constituent to the stream during rain events). This relationship varies and some stream constituent concentrations can even show a combination of both relationships, where there could be an initial dilution of a concentration with increasing baseflow followed by an increase in concentration due to runoff contributing a greater amount of flow and carrying more particulate matter. For example, phosphorus concentrations in the Illinois River in northwest Arkansas decreased as base flow increased, illustrating dilution of point sources, but then concentrations were greater with increased runoff during storm events (Green and Haggard, 2001). Contrarily, nitrate concentrations increased with baseflow, likely due to increased groundwater input, but decreased with surface runoff. The relationship between a constituent concentration and Q will depend on the stream characteristics, the constituent being modelled, and possible sources of the constituent. Thus, multiple methods, both parametric and non-parametric, have been used to adjust concentrations per a given site for exogenous influences in trend analysis (Helsel and Hirsch, 2002; Hirsch et al., 1982).

One method for flow adjustment, which has gained some popularity in the past decade, employs locally weighted regression (LOESS) to adjust concentrations for Q. The residuals from LOESS, or flow-adjusted concentrations (FACs), can then be plotted against time to detect trends (Bekele and McFarland, 2004; Helsel and Hirsch, 2002; White et al., 2004). LOESS can model unknown, even nonlinear, relationships that would otherwise be difficult and sometimes inappropriate for parametric methods (Cleveland, 1979; Jacoby, 2000). LOESS does not make assumptions about the data like parametric methods, so it is useful for adjusting concentrations for flow where the relationship does not necessarily fit a specific theoretical model. But, the smoothing factor (f) must be taken into account for the LOESS regression. The smoothing factor, where $0<f\leq 1$, can increase smoothness with greater f values or follow the variation more closely with lower f values. A past study concluded that, in general, f=0.5 is acceptable for flow-adjustment with LOESS (Bekele and McFarland, 2004). Subsequent studies on trends of several kinds of constituent concentrations at widely varying locations have defaulted to this value in their analyses (e.g., Bekele et al., 2006; Boeder and Chang, 2008; Scott et al., 2011; Wang et al., 2007).

Contrarily, the optimal f ought to be selected for each individual dataset (Cleveland, 1979; Helsel and Hirsch, 2002). The ideal f will produce the smoothest possible regression while still capturing the important curvilinear behavior of the data. Thus, the FACs will better reflect concentrations that are accurately normalized or adjusted for Q. In essence, error from the LOESS regression due to lack of fit should be minimized while predictive power of the regression technique should be maximized. Optimizing LOESS by changing f has been a time-consuming, tedious, and sometimes subjective process in the past, where f was manually changed in an iterative manner until the regression was deemed a good fit for the data (e.g., Trexler and Travis, 1993). The use of automated procedures to select an optimal f is particularly valuable for studies with numerous datasets (e.g., Lettenmaier et al., 1991). Some procedures have calculated the prediction sum of squares (PRESS; Allen, 1974) to find the optimal f value for each analysis (Cleveland, 1979; Lettenmaier et al., 1991). Alternatively, minimizing the bias-corrected Akaike Information Criterion (AIC_c) has been used for selecting f (Bekele and McFarland, 2004).

Recent innovations in software have allowed for more complex and powerful methods of analyzing data, which can be extended to the task of optimizing a LOESS fit (e.g., Lee and Cox, 2010). Here, we propose an automated LOESS optimization procedure for trend analysis of water quality data that uses the "loess.wrapper" function from the "bisoreg" package (Curtis, 2015) in the statistical program R (R Core Team, 2015). This method involves optimizing f through an iterative random K-fold cross-validation process (Kohavi, 1995), thus removing a degree of subjectivity from the flow-adjustment procedure. The objective of this study is to first develop an optimized procedure for flow-adjustment of concentrations for trend analysis. Next, we evaluate monotonic changes in constituent concentrations that have been flow-adjusted using LOESS with the default f=0.5 (sensu White et al., 2004) and with a statistically optimized f. This paper will determine whether f optimization makes a difference in the interpretation of trend results and provide guidance for trend analysis using LOESS to flow-adjust concentrations.

Materials and Methods

Data Set

The water quality data used in this study comes from a monitoring program conducted on two basins in northwest Arkansas, the Upper White River Basin (UWRB) and the Illinois River Watershed (IRW), covering 17 sites. Water samples were collected on a near-weekly basis under base flow conditions and during storm events from July 2009 to June 2015 following quality assurance project plans (QAPPs) approved by the Arkansas Natural Resources Commission (ANRC) and Environmental Protection Agency (EPA). The monitoring sites were located at U.S. Geological Survey (USGS) discharge monitoring stations where mean daily Q was available via the National Water Information System (NWIS; http://waterdata.usgs.gov/nwis) or at stations where the Arkansas Water Resources Center (AWRC) monitors stage continuously to develop a stage-Q rating curve for predicting mean daily Q. The monitored streams vary from small creeks and urban tributaries to large, 5th and 6th order rivers (see Scott et al., 2015). Water samples were analyzed by the AWRC water quality lab (http://arkansas-water-center.uark.edu/waterqualitylab.php) according to standard analytical methods following the QAPP (Scott et al., 2015). The analyzed constituents were: nitrate-nitrogen (NO3-N), total nitrogen (TN), soluble reactive phosphorus (SRP), total phosphorus (TP), total suspended solids (TSS), chloride (Cl), and sulfate (SO4). Thus, there are 119 individual datasets used in the present study. More information about the monitoring sites and analytical methods is available in Scott et al. (2015), which is readily accessible in the AWRC's digital library (see http://arkansas-water-center.uark.edu/publications/msc.php).

General Flow-adjustment Procedure

White et al. (2004) used a three-step trend analysis procedure for water quality data characterized by covariation with Q. First, concentrations and Q were log-transformed to reduce the effect of outliers and produce relatively constant variance in the data (Helsel and Hirsch, 2002; Lettenmaier et al., 1991). Second, the constituent log-concentrations were flow-adjusted via a LOESS fit using f=0.5. Third, the residuals from the LOESS fit, also termed flow-adjusted concentrations (FACs), were analyzed for trends through time. The test for trends can be either parametric (i.e., simple linear regression) or nonparametric, such as a Mann-Kendall or even Seasonal Kendall test (Helsel and Hirsch, 2002). Figure 1 shows this three step process for TSS concentrations at War Eagle Creek using linear regression to determine whether FACs were changing over time.

Optimizing LOESS

The smoothing parameter for LOESS, f, indexes models from nearly a simple linear regression (f = 1) to approximately an n^{th}-order polynomial at low values (Jacoby, 2000). Overfitting the model can lead to unwanted variance, or "wiggle", in the fit while underfitting can create a biased model that fails to capture the important variance in the data. This concept, known as the bias-variance tradeoff, is well-known in the field of machine learning where the true relationship being modelled must be inferred from the data (Hastie et al., 2009). Characterizing the bias-variance tradeoff requires an estimate of prediction error since the true prediction error for some model cannot be known. Estimates of model prediction error can help identify an ideal amount of model complexity, balancing bias and variance. When estimating prediction error, using all the available data is not appropriate since increasing model complexity can only improve model performance. In order to make sure a model generalizes well (i.e., minimizes prediction error), validation methods can be used to estimate prediction error for a given model (Hastie et al., 2009).

Cross-validation

To estimate prediction error, several model validation techniques can be employed (e.g., Moriasi et al., 2012). One robust method is *K*-fold cross-validation (Kohavi, 1995). *K*-fold cross-validation (*k*-fold CV) randomly partitions data into *K* number of subsets (folds), fits a model to *K*-1 of the folds, and uses the fold that was left out as the test set. This is repeated for each fold so that each observation is used for validation exactly one time while being used for training the model *K*-1 times. In the case where prediction mean squared error (MSE) is being estimated for a given *f*, the CV procedure is given by:

$$CV(\widehat{m}\,, f) = \frac{1}{N}\sum_{i=1}^{N} L(y_i, \widehat{m}^{-k(i)}(x_i, f)) \qquad (1)$$

where $\widehat{m}$ is the fitted LOESS function with the given *f* parameter, *N* is the number of observations, y_i is the *i*th observed value, *k* is an indexing vector {1,…*K*} that randomly partitions the observations, x_i is the *i*th predictor variable and *L* is the squared error loss function given by:

$$L\big(y, f(x)\big) = (y - f(x))^2 \qquad (2)$$

The smoothing parameter is tuned so that the estimated MSE from equation (1) is minimized. In our study, we compared a sequence of *f* values from 0.1 to 1 indexed by a value of 0.05.

Lettenmaier et al. (1991) used the PRESS statistic (Allen, 1974) as a method to determine an optimal *f* which is actually *K*-fold CV with *K* equal to *N* (also known as leave-one-out CV). However, Kohavi (1995) points out that too large of a *K*, such as in leave-one-out CV, results in poor estimations of prediction error due to high variance (i.e., largely dependent on the sample) of the error estimation. Contrarily, too small of a K (e.g., 2-fold CV) can be a pessimistic, or biased, estimation of prediction error.

Since K-fold CV can never perform perfectly no matter the choice of K (Bengio and Grandvalet, 2004), heuristic rules on how to apply K-fold CV are developed for different applications such as a 5 x 2 CV test (5 iterations of 2-fold CV t-test) in the context of classification learning algorithms (Dietterich, 1998). In the present study, we compare the variability in choice of f that minimizes the prediction error estimated by K-fold CV from $K = 5$ to $K = 45$ with the statistics environment R (R Core Team, 2015). The 'loess.wrapper' function from the bisoreg package (Curtis, 2015; CRAN: bisoreg) effectively incorporates K-fold CV with fitting LOESS models and was modified as needed to generate additional output (e.g., MSE estimates). All 7 constituents at one study site, West Fork of the White River, were analyzed via the flow-adjustment procedure to find the optimized f (f_{opt} ; the median value of the K-fold CV selections) for each value of K for 1000 iterations. This information was used to guide our selection of K.

An appropriate number of iterations of K-fold CV were needed to consistently select f_{opt}. Here, we seek to use a minimal number of iterations to minimize computation costs. To compare the number of iterations of K-fold CV necessary to select the median f_{opt} in a consistent manner, we conducted the flow-adjustment procedure with a relatively small value of K for an arbitrary small number of iterations and for 1000 iterations for all 119 datasets. We plotted the f_{opt} selected for $K = 10$ with iterations of 10 and 1000 to evaluate apparent differences relative to the 1:1 line across streams and constituents.

The last step in developing our heuristic K-fold CV flow-adjustment procedure was to recommend a final K. The choice of K was analyzed by comparing variability in f_{opt} with our selection of K and a relatively large value of K = 35. For this simulation, the flow-adjustment procedure with both values of K was conducted for all 119 datasets for only 10 iterations. These selected f_{opt} were plotted against each other to observe differences relative to the 1:1 line.

Trend Comparisons

We used the three step process to evaluate monotonic trends in constituent concentrations, where flow-adjustment was performed using the standard f=0.5 ($f_{0.5}$) and f_{opt} using our rule for K-fold CV; this approach was used on all 119 datasets. The monotonic trends were evaluated using simple linear regression on the plot of FACs over time, providing an estimate of the trend magnitude (interpreted as percent change per year in constituent concentration; % change year^{-1}) as well as significance behind the trend from the overall F-test (Helsel and Hirsch, 2002). The equation for determining % change year^{-1} is given by:

$$\%\, change\ year^{-1} = (e^{m} - 1) \times 100 \tag{3}$$

where m is the linear regression slope with the necessary time unit conversions.

All values of % change year^{-1} in constituent concentrations deemed significant at α=0.10 were compared between flow-adjustment procedures using $f_{0.5}$ and f_{opt}. It must be noted here that we assume that trends in the water quality data are monotonic and that the stream systems exhibit stationarity, such that we can determine whether optimizing f has an influence on trend interpretation across these sites and constituents (n=119 datasets).

Results and Discussion

Developing a heuristic **K-***fold CV rule for trend analysis*

The simulation of the *K*-fold CV procedure for flow-adjustment at West Fork of the White River showed variability with f_{opt} (Figure 2), providing two key observations. First, the choice of f_{opt} varies depending on which constituent is being analyzed at this site. NO_3, TN, SRP and Cl concentrations were better characterized with relatively larger values for f ($f > 0.6$) while TP and TSS concentrations were fit better with a smaller f value ($f < 0.3$). The median f_{opt} was not 0.5 with an increasing value of *K* and 1000 iterations for this site across these constituents. Second, selection variability tends to be reduced after K=10 across most of these constituents resulting in a similar median f_{opt} for a given constituent with increasing *K*. This reduced variability in f_{opt} agrees with observations on the performance of *K*-fold CV in other settings (Arlot and Lerasle, 2015; Kohavi, 1995). In a theoretical approach, Arlot and Lerasle (2015) evaluated the choice of *K* (notated as *V* in their paper) in a least-squares density estimation application and showed that the error estimation with a small *K* is improved greatly when *K* is increased to K=5 or 10, but improvements diminish quickly with greater *K* values. Kohavi (1995) used multiple well-known datasets (e.g., Fisher's iris dataset) to illustrate the performance of classification algorithms in terms of accuracy estimated via *K*-fold CV. Kohavi (1995) noted that the choice of *K* affects the stability of the learning algorithm with small values (e.g., K=2) generating pessimistically biased error estimates, especially with small datasets, while larger *K* (K=20) can be more variable in estimating model prediction error. Given our empirical study of *K*-fold CV at West Fork of the White River and the insight from statistical literature, we opted to further analyze the performance of K=10.

Since the K-fold CV procedure has inherent randomness due to the fold partitioning procedure, multiple iterations can be helpful to account for the variability in the error estimate of K-fold CV (Kohavi, 1995). To compare the effect that the number of iterations has on the consistency of the CV procedure with K=10, the selected f_{opt} for 10 x 10 CV (10 iterations of 10-fold CV) and for 1000 x 10 CV for all datasets are shown in Figure 3. Generally, the selected f_{opt} values follow a 1:1 relationship, where most f_{opt} values for 10 and 1000 iterations were within ±0.1. The exception, which deviated the most, was the flow-adjustment of Cl concentrations at the Kings River, where the median f_{opt} varied from 0.35 with 1000 iterations to 0.875 with 10 iterations.

The comparison in the selected f_{opt} for 10 and 35 folds with 10 iterations for all datasets is shown in Figure 4. The two choices of K show a good amount of agreement on selecting f_{opt} with the majority of points differing in choice of f_{opt} by less than ±0.1. Only a few cases resulted in widely different values for f_{opt}. For example, the largest differences in observed median f_{opt} were Cl concentrations at Kings River (0.3 for K=35 folds, 0.875 for K=10 folds), Cl concentrations at WR45 (0.2 for K=35 folds, 0.775 for K=10 folds), and TSS concentrations at Spring Creek (0.1 for K=35 folds, 0.725 for K=10 folds).

The points that deviate the most from the 1:1 line in figures 3 and 4 are the result of selecting f_{opt} to minimize for very small differences in MSE of the LOESS fit. For example, Figure 5 shows the test error relationship with choice of f for each constituent at the Kings River using 10 x 10 CV as well as the conceptual test/training error relationship in machine learning applications described by Hastie et al. (2009) for comparison. In the Cl subplot (Figure 5), the estimated MSE is near its minimum for f=0.3 as well as f=0.85. In this case, the discrepancy in the LOESS fit due to choice of either f is negligible. Thus, the choice of f_{opt} given by 10 x 10 CV performs well in minimizing prediction error for the present application.

We prefer to use the lesser value of K=10 in order to cut computation costs, reduce the risk of employing a potentially variable (i.e., dependent on the specific sample data being used) estimation of prediction error via a larger K (Kohavi, 1995), and to avoid the statistical issue of correlated training sets (Dietterich, 1998). K values ranging from 5 to 10 are often used in statistical learning frameworks including hydrological studies such as evaluating model performance when predicting groundwater NO_3 concentrations (10-fold CV; Nolan et al., 2015); selecting kernel function parameters for modelling flood-prone areas of Malaysia (5-fold CV; Tehrany et al., 2014) and for predicting streamflow (5-fold CV; He et al., 2014); and evaluating regression tree models for reservoir algal growth responses to physical and nutrient content variables (10-fold CV; Park et al., 2015). Our empirical analysis further supports the use of K=10 for cross-validating various hydrological models, including flow-adjusting concentrations for trend analysis.

It is interesting to note that the error relationships given in Figure 5 generally reflect the pattern of the test error shown in the conceptual figure (see top left plot). However, the shape of this relationship is different for not only each constituent dataset at the Kings River, but is unique to each dataset at all sites in our study. This can be related to the study site, the constituent being analyzed, or even the number of observations in the dataset (e.g., Warrick et al., 2013). When flow-adjusting suspended sediment concentrations from six sites in California using the general procedure described above, Warrick et al. (2013) selected f manually so that LOESS followed the curvature in the data. Warrick et al. (2013) noted that smaller datasets benefitted from relatively larger f values and vice versa for larger datasets, although their choice of f only ranged from 0.1 to 0.2. In comparison, the f_{opt} values in this study varied from 0.25 to 0.9 for TSS (analogous to suspended sediment concentrations), and from 0.1 to 1.0 across all constituents.

Applying 10 x 10 CV rule for trend analysis

With the selected 10 x 10 CV rule, all 119 datasets were flow-adjusted for trend analysis. Figure 6 shows the distribution of selected f_{opt} values across all datasets and for each constituent. Overall, f_{opt} spans the full range of possible f values examined and tends to occur most frequently near 0.6 with more extreme values (i.e., f_{opt} = 0.1 or 1.0) occurring less frequently. On a constituent basis, the f_{opt} selections showed near-uniform distributions for SRP and Cl (Figure 6, subplots D and G, respectively) while selected f_{opt} values in TP and SO_4 (subplots E and H, respectively) resembled the overall distribution. For these datasets, there appears to be no pattern between the type of constituent (e.g., dissolved versus particulate forms) and the choice of f, reflecting a case-specific need for determining f_{opt}.

Given the tendency of most datasets in our analysis to have a mid-range f_{opt} (Figure 6), the value ($f_{0.5}$) recommended by Bekele and McFarland (2004) likely works well for flow-adjusting concentrations when analyzing water quality trends. Indeed, when applying both f_{opt} and $f_{0.5}$ to all our datasets and determining the % change year^{-1} in concentrations using the slope from simple linear regression (α=0.10), there is no discernible difference in interpretation of trend magnitude (Figure 7). Additionally, out of 119 datasets, there were only 3 datasets where the trend was statistically significant using f_{opt} but not significant using $f_{0.5}$ at the 10% significance level (data not shown). There was only one instance of the opposite case where NO_3 concentrations at Spring Creek showed a statistically significant trend with $f_{0.5}$ (p=0.06) but no significant trend with f_{opt} = 1.0 (p = 0.18). However, this site had a shorter record of data (3.5 years) and trend analysis techniques may be less reliable for data with insufficient time spans.

The use of non-parametric trend detection techniques may be more appropriate than simple linear regression for water quality data (see discussion by Esterby, 1996 and Hirsch et al., 1991).

The FACs generated by the proposed technique here can be analyzed for trends via methods such as the Seasonal Kendall test just as easily as simple linear regression (Hirsch et al. 1991). Additional studies could evaluate the efficacy of the method on other datasets that vary in site characteristics, the nature of the constituent being analyzed, and data availability. The sites used in this study were relatively data rich, where near weekly concentration data were available. It would be interesting to determine if using f_{opt} has an influence on trend interpretations for sites with less frequent sampling (e.g., monthly or bimonthly data), which is common in water quality studies. More complex work in understanding stream biogeochemical processes could potentially benefit from the proposed flow-adjustment procedure where it is necessary to include the constituent's covariation with Q in a model, such as in the time-series model with autoregressive moving average (ARMA) error developed by Abaurrea et al. (2011).

Conclusion

Trend analyses of stream water quality should consider carefully how to model the covariation of a constituent with Q in order to produce accurate estimates of trends. Here, we build upon the flow-adjustment procedure using LOESS so that the smoothing parameter, and thus the LOESS fit, is tailored to each dataset. We recommend using 10 x 10 CV to determine the optimal f value to use. Though we found that the interpretation of the monotonic trend magnitude and direction was not different using either f_{opt} or $f_{0.5}$, we cannot speak to what effect there may be when f_{opt} is applied to other sites containing more/less data, differently behaving constituents, or other potential factors. However, the 10 x 10 CV is a robust method to determine the optimal fit of LOESS to water quality data and should perform well in other environments.

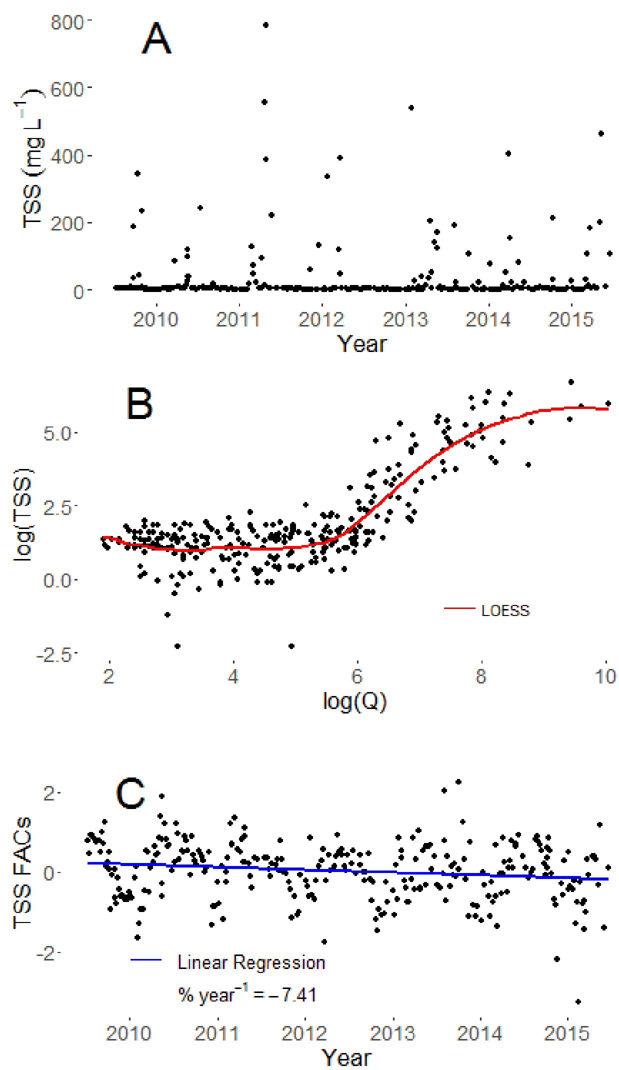

Figure 1. Three step trend process with (A) total suspended solids (TSS) concentrations at War Eagle Creek, AR from 2009 to 2015, (B) log-transformed TSS concentrations and log-transformed streamflow (Q) with a LOESS fit (f=0.5), and (C) residuals from the LOESS fit, or flow-adjusted concentrations (FACs), plotted as a function of time with a linear regression as the trend test ($p < 0.01$) and the trend slope given in percent change in TSS per year.

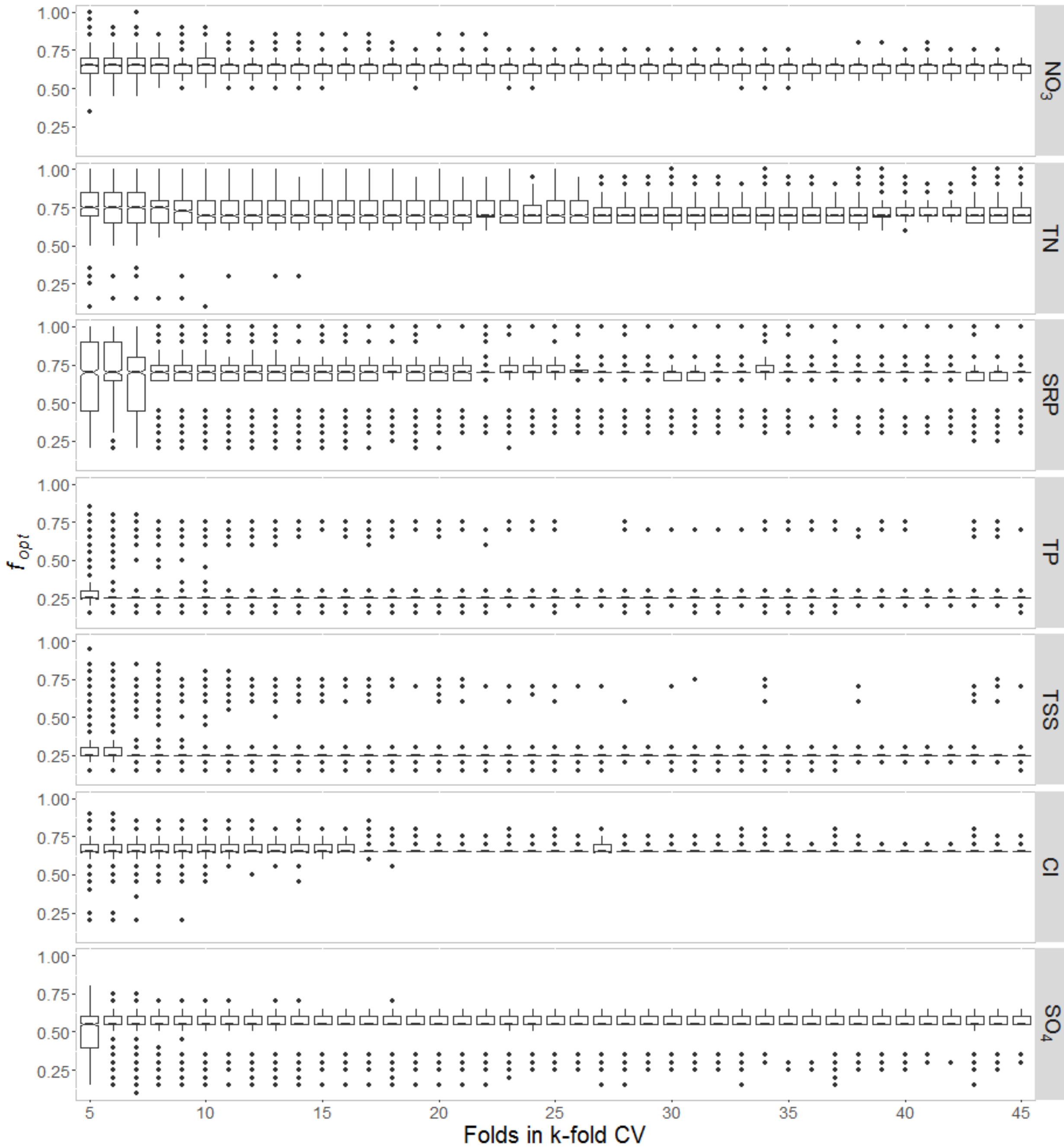

Figure 2. Boxplots of the optimal sampling proportion (f_{opt}) chosen for a given value of K in 1000 iterations of K-fold cross-validation for 7 constituents at West Fork of the White River.

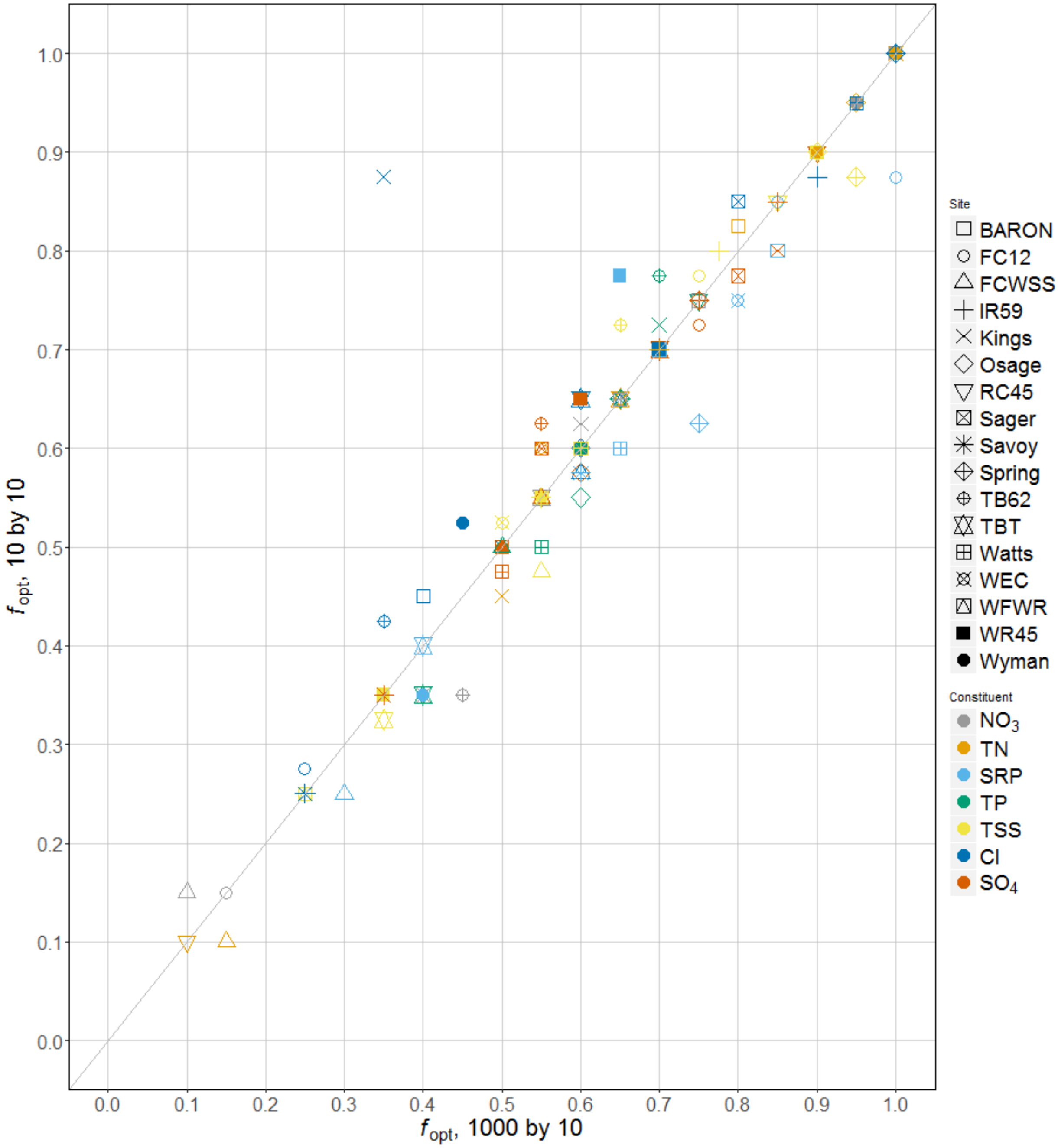

Figure 3. Selection of optimal sampling proportion (f_{opt}) for all datasets when using $K = 10$ for 10 and 1000 iterations of K-fold cross-validation (10 by 10 and 1000 by 10, respectively).

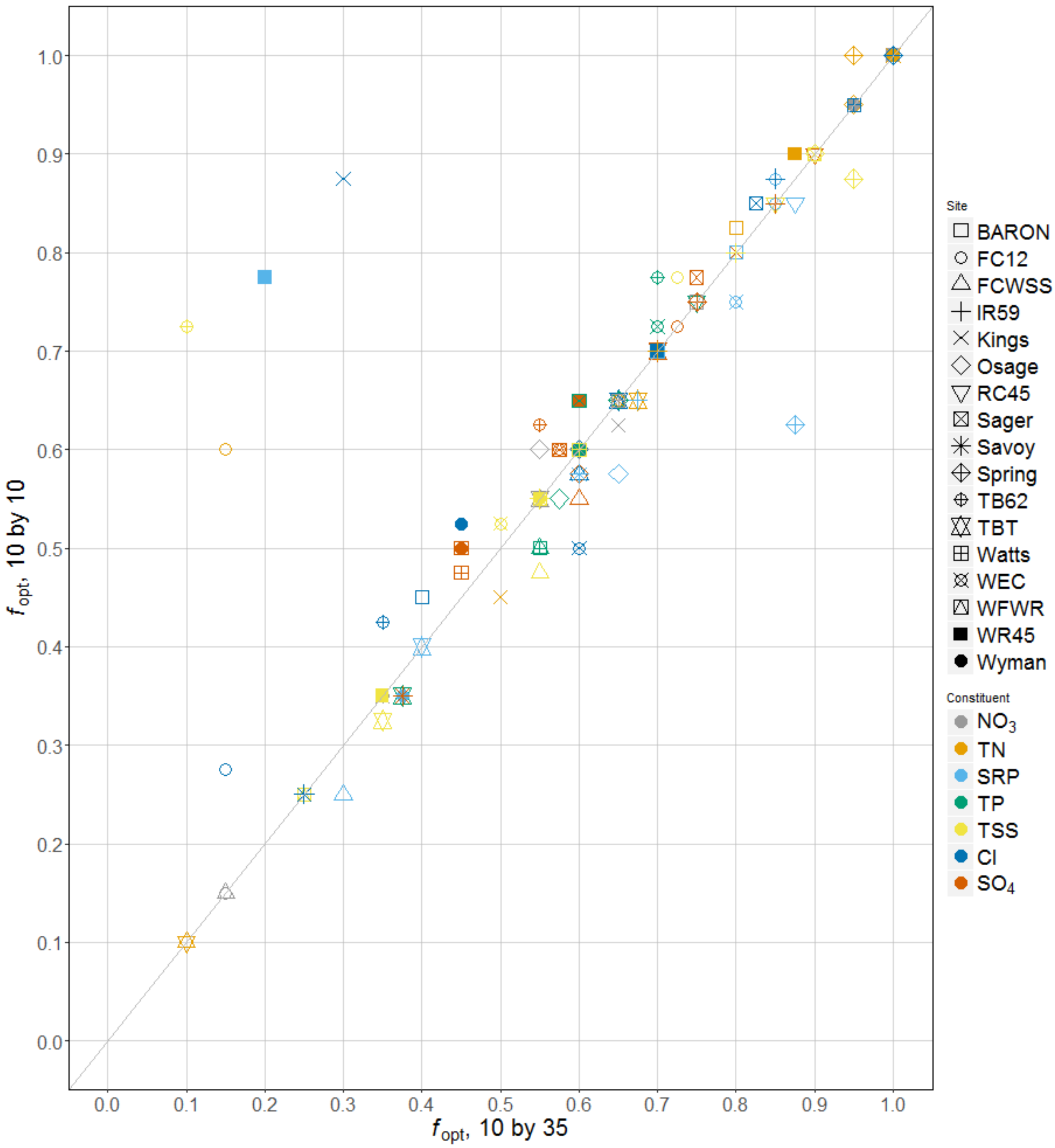

Figure 4. Selection of optimal sampling proportion (f_{opt}) for all datasets when using K = 10 and K = 35 folds for 10 iterations of K-fold cross-validation (10 by 10 and 10 by 35, respectively).

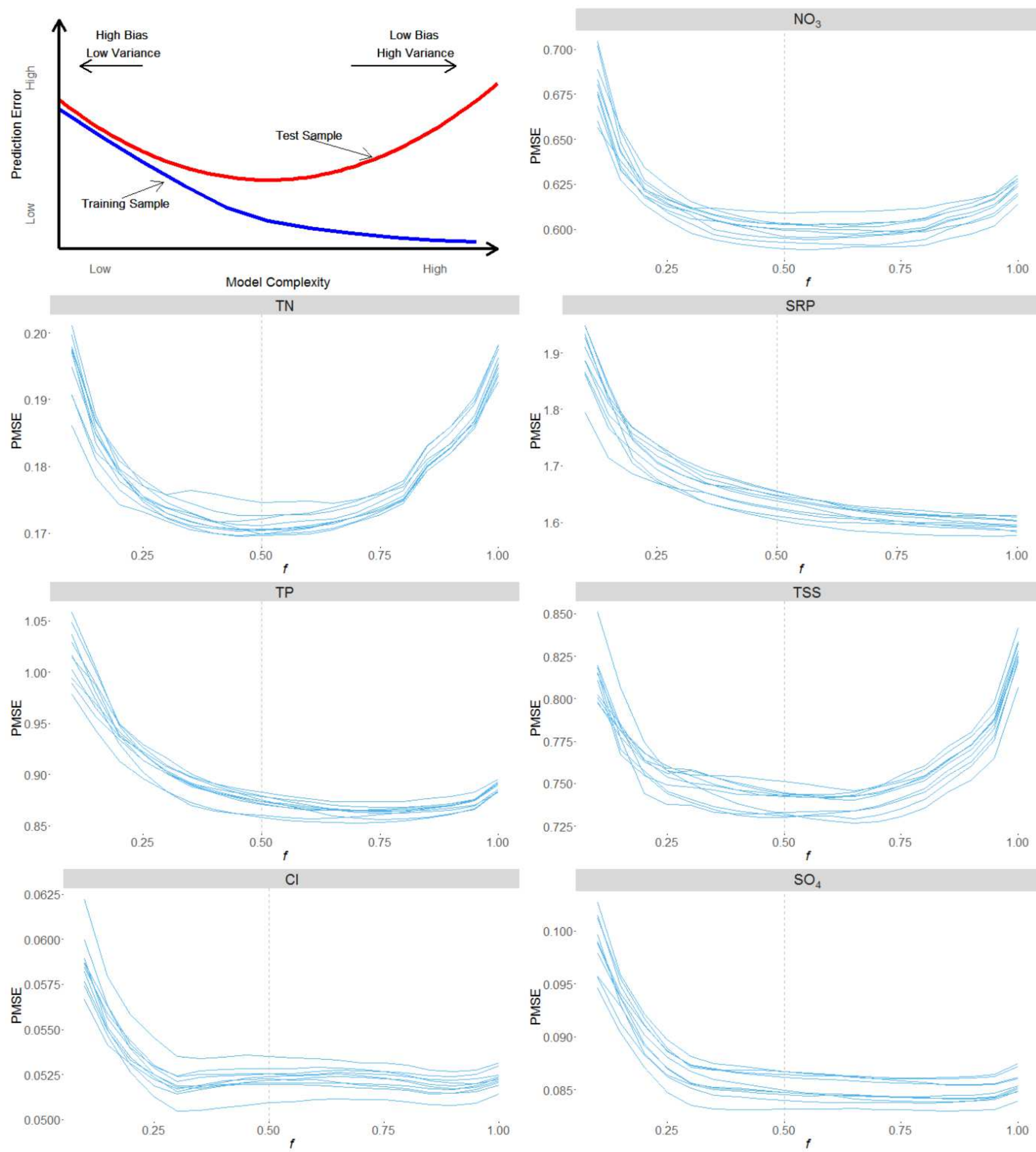

Figure 5. Test error plotted against model complexity. The top left plot is a conceptual figure adapted from Hastie et al. (2009), where training sample refers to using all available data for estimating prediction error rather than a subset (test sample). For the constituent plots, estimated mean squared error of prediction (PMSE) via 10-fold cross-validation across a *f* values ranging from 0.1 to 1.0 (note that larger *f* gives a simpler LOESS model) is shown for the datasets at the Kings River. Each line represents an iteration of 10-fold cross-validation. Note that the y-scale for each subplot is different due to differing magnitudes in the underlying constituent concentrations (PMSE is estimated from the log-transformed data). The dashed vertical line is a reference for $f = 0.5$.

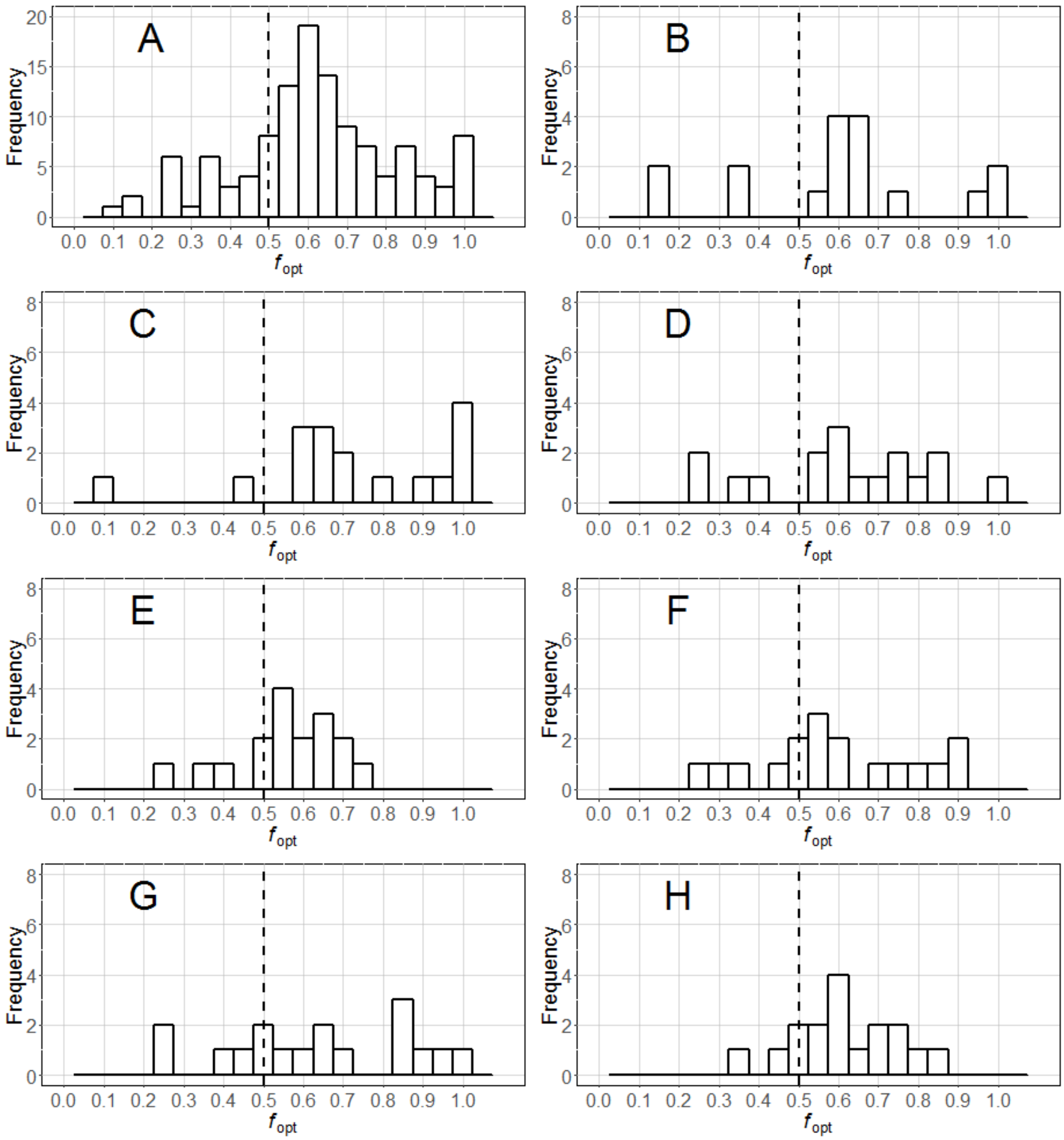

Figure 6. Distribution of f_{opt} selected via 10 x 10 cross-validation for all 119 datasets (A), and the subsets for nitrate-nitrogen (B), total nitrogen (C), soluble reactive phosphorus (D), total phosphorus (E), total suspended solids (F), chloride (G), and sulfate (H). The dashed vertical line is a reference for f=0.5.

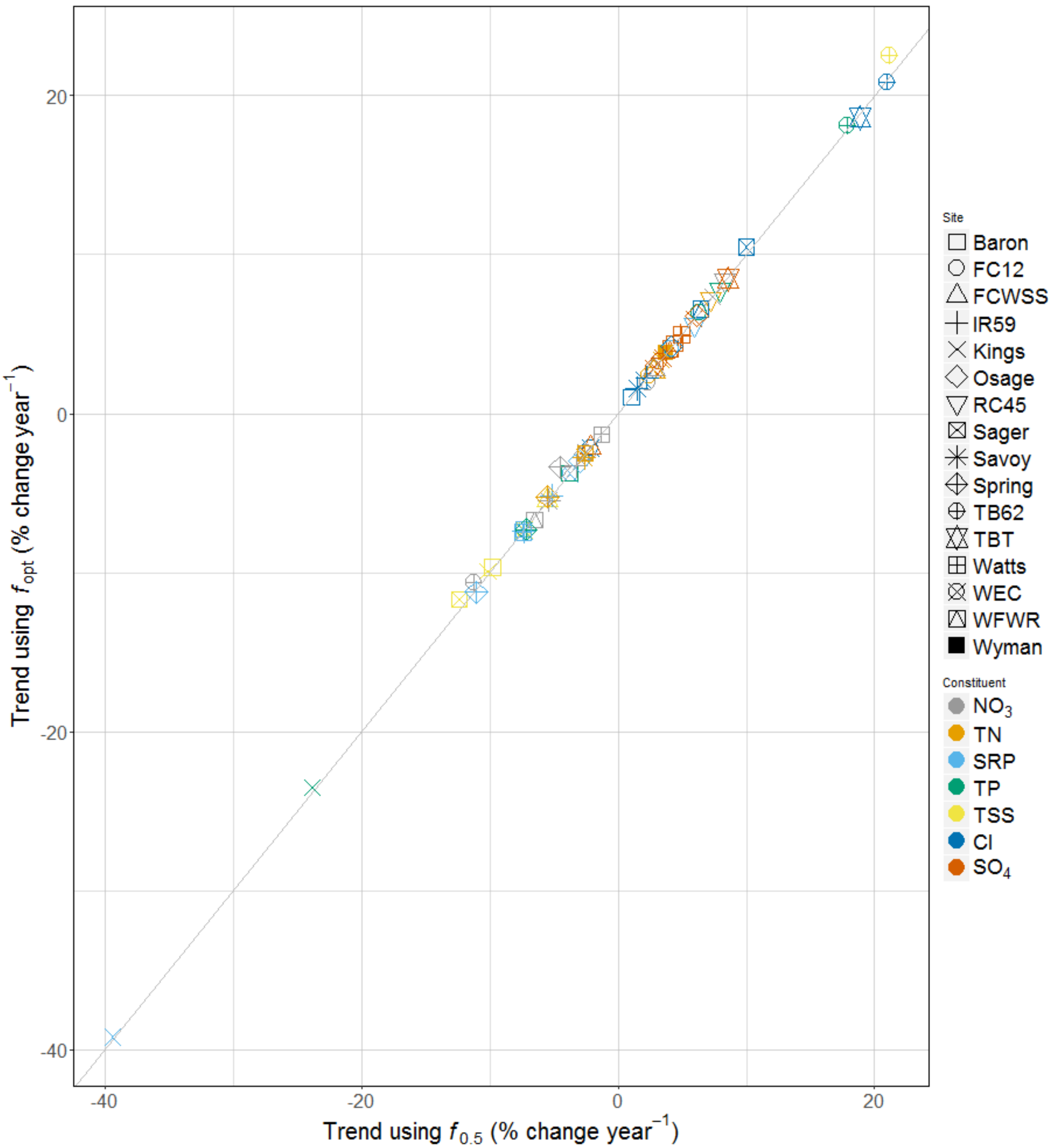

Figure 7. Monotonic trend magnitudes given as percent change in concentration per year (% change year^{-1}) for all datasets using a sampling proportion of 0.5 ($f_{0.5}$) and an optimal sampling proportion (f_{opt}) in the flow-adjustment procedure. Only trends that are significant at the 10% level are shown.

Works Cited

Abaurrea, J., J. Asín, A.C. Cebrián, and M.A. García-Vera. 2011. Trend analysis of water quality series based on regression models with correlated errors. Journal of Hydrology 400 (3-4):341–352.

Allen, D.M. 1974. The relationship between variable selection and data augmentation and a method for prediction. Technometrics 16(1):125-127.

Arlot, S. and M. Lerasle. 2016. Choice of V for V-fold cross-validation in least-squares density estimation. To appear. Journal of Machine Learning Research (17). Accessed at https://hal.archives-ouvertes.fr/hal-00743931/en/

Bekele, A. and A. McFarland. 2004. Regression-based flow adjustment procedures for trend analysis of water quality data. Transactions of the American Society of Agricultural Engineers 47(4):1093-1104.

Bekele, A., A.M.S. McFarland, and A.J. Whisenant. 2006. Impacts of a manure composting program on stream water quality. Transactions of the American Society of Agricultural and Biological Engineers 49(2):389-400.

Bengio, Y. and Y. Grandvalet. 2004. No unbiased estimator of the variance of k-fold cross-validation. Journal of Machine Learning Research 5: 1089-1105.

Boeder, M. and H. Chang. 2008. Multi-scale analysis of oxygen demand trends in an urbanizing Oregon watershed, USA. Journal of Environmental Management 87:567-581.

Bronster, A. and G. Bürger. 2002. Effects of climate and land-use change on storm runoff generation: present knowledge and modeling capabilities. Hydrologic Processes 16: 509-529.

Cleveland, W.S. 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 74(368): 829-836.

Curtis, S. McKay Curtis (2015). bisoreg: Bayesian Isotonic Regression with Bernstein Polynomials. R package version 1.4. http://CRAN.R-project.org/package=bisoreg

Dietterich, T.G. 1998. Approximate statistical tests for comparing supervised classification learning algorithms. Neural Computation 10: 1895-1923.

Esterby, S.R. 1996. Review of methods for the detection and estimation of trends with emphasis on water quality applications. Hydrological Processes 10: 127-149.

Green, W.R. and B.E. Haggard. 2001. Phosphorus and nitrogen concentrations and loads at Illinois River south of Siloam Springs, Arkansas, 1997-1999. United States Geological Survey. Water-Resources Investigations Report 01-4217.

Haggard, B.E. 2010. Phosphorus concentrations, loads, and sources within the Illinois River drainage area, northwest Arkansas, 1997-2008. Journal of Environmental Quality 39: 2113-2120.

Harmel, R.D., K.W. King, B.E. Haggard, D.G. Wren, J.M. Sheridan. 2006. Practical guidance for discharge and water quality data collection on small watersheds. Transactions of the American Society of Agricultural and Biological Engineers 49(4): 937-948.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. Overview of supervised learning. In The Elements of Statistical Learning: Data Mining, Inference, and Prediction, (pp. 9-42). Germany: Springer series in statistics. Accessed at http://web.stanford.edu/~hastie/pub.htm

He, Z., X. Wen, H. Liu, and J. Du. 2014. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. Journal of Hydrology 509: 379-386.

Helsel, D.R., and R.M. Hirsch. 2002. Statistical methods in water resources. In Hydrologic Analysis and Interpretation, (pp. 1-510). United States Geological Survey.

Hirsch, R.M., J.R. Slack, and R.A. Smith. 1982. Techniques of trend analysis for monthly water quality data. Water Resources Research 18(1): 107-121.

Hirsch, R.M., R.B. Alexander, and R.A. Smith. 1991. Selection of methods for the detection and estimation of trends in water quality. Water Resources Research 27(5): 803-813.

Jacoby, W.G. 2000. Loess: a nonparametric, graphical tool for depicting relationships between variables. Electoral Studies 19: 577-613.

Kohavi, R. 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. International Joint Conference on Artificial Intelligence 14(2): 1137-1145.

Lee, J.S. and D.D. Cox. 2010. Robust smoothing: smoothing parameter selection and applications to fluorescence spectroscopy. Computational Statistics and Data Analysis 54(12): 3131-3143.

Leh, M., S. Bajwa, and I. Chaubey. 2011. Impact of land use change on erosion risk: an integrated remote sensing, geographic information system and modeling methodology. Land Degradation and Development 24(5): 409-421.

Lettenmaier, D.P., E.R. Hooper, C. Wagoner, and K.B. Faris. 1991. Trends in stream quality in the continental United States, 1978-1987. Water Resources Research 27(3): 327-339.

Miltner, R.J. 2015. Measuring the contribution of agricultural conservation practices to observed trends and recent conditions in water quality indicators in Ohio, USA. Journal of Environmental Quality 44: 1821-1831.

Moriasi, D.N., B.N. Wilson, K.R. Douglas-Mankin, J.G. Arnold, and P.H. Gowda. 2012. Hydrologic and water quality models: use, calibration, and validation. Transactions of the American Society of Agricultural Engineers 55(4): 1241-1247.

Nolan, B.T., M.N. Fienen, D.L. Lorenz. 2015. A statistical learning framework for groundwater nitrate models of the Central Valley, California, USA. Journal of Hydrology 531: 902-911.

Park, Y., Y.A. Pachepsky, K.H. Cho, D.J. Jeon, and J.H Kim. 2015. Stressor-response modeling using the 2D water quality model and regression trees to predict chlorophyll-a in a reservoir system. Journal of Hydrology 529: 805-815.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for statistical Computing, Vienna, Austria. http://www.R-project.org/.

Scott, E.E., Z.P. Simpson, and B.E. Haggard. 2015. Constituent loads and trends in the Upper Illinois River Watershed and Upper White River Basin. Arkansas Water Resources Center, Fayetteville, Arkansas, MSC Publication 377: 89 pp.

Scott, J.T., B.E. Haggard, A.N. Sharpley, and J.J. Romeis. 2011. Change point analysis of phosphorus trends in the Illinois River (Oklahoma) demonstrates the effects of watershed management. Journal of Environmental Quality 40:1249-1256.

Trexler, J.C. and J. Travis. 1993. Nontraditional regression analyses. Ecology 74(6):1629-1637.

Wang, X., A.M. Melesse, M.E. McClain, and W. Yang. 2007. Water quality changes as a result of coalbed methane development in a Rocky Mountain watershed. Journal of the American Water Resources Association 43(5):1383-1399.

Warrick, J.A., M.A. Madej, M.A. Goñi, and R.A. Wheatcroft. 2013. Trends in the suspended-sediment yields of coastal rivers of northern California, 1955-2010. Journal of Hydrology 489: 108-123.

White, K.L., B.E. Haggard, and I. Chaubey. 2004. Water quality at the Buffalo National River, Arkansas, 1991-2001. Transactions of the American Society of Agricultural Engineers 47(2): 407-417.

II. Are Phosphorus Concentrations Still Declining in the Illinois River Watershed, Northwest Arkansas?

Abstract

Phosphorus (P) management at the watershed-scale has the ultimate goal of improving water quality. A case study would be the Illinois River Watershed (IRW), located in northwest Arkansas and Oklahoma, with its history of P enrichment from land-applied poultry litter and effluent discharges, where reduced effluent P concentration resulted in significant reduction in total P (TP) concentrations in the Illinois River from 2002 to 2008. However, it is unknown whether the efforts taken to further reduce effluent P, as well as management of land-applied poultry litter, have continued to decrease P concentrations in the Illinois River. In this study, we examine three streams for changes in flow-adjusted P concentrations from 2009 to 2015. Phosphorus, both soluble reactive P (SRP) and TP), decreased 12 and 9% year^{-1}, respectively, at the most upstream site, Spring Creek, which was the site in closest proximity to a major effluent discharge. However, only SRP showed significant decrease further downstream at Osage Creek (3% year^{-1}) and the Illinois River (6% year^{-1}). These decreases likely reflect the decreases in effluent P during the study period (11% year^{-1} decrease in monthly average TP concentrations across all major wastewater facilities). This study highlights the importance of managing effluent inputs when seeking water quality improvements, while possible benefits from landscape management likely exhibit a lag time in taking effect.

Introduction

Water quality issues have forced changes in regulations and nutrient management, resulting from lawsuits, watershed planning, total maximum daily loads (TMDLs) implementation, and other activities. The trans-boundary Illinois River Watershed (IRW) in Arkansas and Oklahoma is a perfect example, where Oklahoma established a scenic rivers total phosphorus (TP) criterion (0.037 mg L^{-1}; OWRB, 2002), prompting the Oklahoma Attorney General to file a lawsuit against members of the poultry industry in northwest Arkansas. The states signed two statements of joint principles and actions (2003 and 2013), and several new regulations were put into place, as well as a joint study to evaluate algal response to increasing P concentrations. The aim of these jointly signed statements and new regulations was to reduce both non-point and point sources of P, while determining an appropriate P threshold.

The primary focus on non-point sources in the IRW has been the poultry industry and the management of the resultant poultry litter. Poultry litter represents two P sources in the watershed, the legacy of soil P from historic applications and that applied to the landscape - these sources were addressed through new programs and policies in Arkansas and Oklahoma. Poultry litter applications to fields within the IRW must conform to state nutrient management (i.e., the P index; Sharpley et al., 2003). The subsequent reduction in litter application coincided with the development of a litter export program. Established in the mid-2000s, the program now annually exports around 90,000 metric tons of poultry litter to neighboring watersheds, thus removing a significant portion of the 250,000 metric tons produced within the IRW on average (Herron et al., 2012). However, these changes in management might not result in a water quality response (i.e., decreasing stream P concentrations) for decades (e.g., Meals et al., 2010; Sharpley et al., 2013).

Another scrutinized source of P in the IRW has been the municipal wastewater treatment plants (WWTPs). After voluntary reductions in effluent TP concentrations at two major WWTPs in the IRW, the proportion of the TP load from WWTPs decreased from 40% to less than 15% by 2006

(Haggard, 2010). This resulted in a significant decrease in TP concentrations in the Illinois River at the Arkansas-Oklahoma border from 2003 to 2009 (Scott et al., 2011). Recently, regulations specific to the IRW have required all WWTPs to meet a TP limit of 1 mg L^{-1} (effective 2012; APCEC, 2015) and one major WWTP (i.e., Springdale's WWTP) in northwest Arkansas has changed management to further reduce effluent P inputs.

We hypothesize that further reductions in effluent TP should be reflected in downstream P concentrations within the IRW, thus continuing the decreases previously observed within the Oklahoma portion of the watershed (Scott et al., 2011). In this study, we analyze soluble reactive P (SRP) and TP concentrations in the IRW from 2009 to 2015 at three monitoring sites downstream from major WWTP discharges. The concentration data were separated according to baseflow contribution and subsequently analyzed for trends in flow-adjusted concentrations (FACs).

Methods

Site Description

The IRW (HUC 11110103) lies in northwest Arkansas, originating southwest of Fayetteville, Arkansas, and flows into northeastern Oklahoma before emptying into the Arkansas River. IRW drains approximately 432,823 ha, with 54% of the drainage on the Oklahoma side (OCC, 2010) and is characterized by the Ozark Highlands and the Boston Mountains Ecoregions. The Arkansas side of the watershed is mostly pasture (45%), forest (37%), and urban (13%) (CAST, 2006), and the IRW has seen a large increase in population from 1990 to 2000 (131,000 to 194,000). The population has continued to increase by over 30% (2010 Census Data) in the Arkansas portion since 2000.

Water Quality Sampling

Water quality samples were collected at USGS gaging stations at Spring Creek (USGS 07194933), Osage Creek (USGS 07195000), and Illinois River at Arkansas Highway 59 (Illinois River; USGS 07195430), shown in Figure 1. Grab samples were collected from the stream's centroid of flow on a near weekly basis including both baseflow and storm events. Samples were analyzed at the Arkansas Water Resources Center (AWRC) certified water quality lab using approved standard methods for the analysis of water samples (AWRC, 2016). Analytes included nitrate-nitrogen, total nitrogen, SRP, TP, total suspended solids (TSS), chloride, and sulfate (see Scott et al., 2015); however, we focused on SRP, TP and TSS. Osage Creek and Illinois River data spanned July 2009 through June 2015 (6 years), while the record for Spring Creek was only from January 2012 to June 2015 (~3.5 years). Discharge data (cfs) were available from the USGS National Water Information System (NWIS) for each site across the study period.

All of these sites are downstream from major WWTPs in northwest Arkansas (Figure 1). The WWTP in Springdale discharges into Spring Creek about 1.8 km upstream from our sampling site. The sampling site on Osage Creek is downstream from the effluent discharge of the Springdale and Rogers WWTPs, as well as a rural treatment facility built by the Northwest Arkansas Conservation Authority (NACA). The sampling site on the Illinois River is influenced by all the above mentioned WWTPs as well as the Fayetteville Westside treatment facility. Monthly averaged effluent TP concentrations were provided by these WWTPs and are shown in Figure 2.

Data Analysis

Daily mean discharge (Q) was used to generate the baseflow discharge (BQ) record via hydrograph separation (Eckhardt, 2005). This method involves a procedure developed by Nathan and McMahon (1990) with an improvement on the BQ filter algorithm (Chapman, 1991), which produces hydrologically reasonable results (Eckhardt, 2008). Filter values recommended for perennial streams with porous aquifers were used (Nathan and McMahon, 1990; Eckhardt, 2005), since the IRW has underlying karst geology (Jarvie et al., 2014). An example hydrograph for the Illinois River during May 2011 is provided in Figure 3, illustrating how the separation technique characterizes the flashy peaks as mostly surface runoff while the BQ has a dampened, smoothed response.

To define which data points in our analysis were representative of storm or base conditions, a baseflow fraction (BFF; BQ/Q) was generated for each sample date. Considering that these streams are normally relatively clear under base conditions (e.g., 10 NTU is the water quality standard (Arkansas Regulation 2; APCEC, 2015), and turbidity during base flow at Illinois River is ~5 NTU on average (data not shown)), we expect that total suspended solids (TSS) concentrations during storm events should be characteristically different from baseflow conditions. So, a changepoint with 95% confidence intervals in the relationship between log-transformed TSS concentrations and BFF was identified at all three sites using a nonparametric changepoint analysis technique (NCPA; see Qian et al., 2003). Since the changepoints, ranging from 0.42 for Spring Creek to 0.59 for Illinois River, had overlapping confidence intervals, a common value of 0.6 was used as the changepoint across all three sites (Figure 4). Water quality data with BFF of 0.6 or greater (i.e., 60% or more of Q is BQ) were classified as baseflow data, while data with a lesser BFF were termed as storm data. This is similar to the 70% value used in White et al. (2004) for separating water quality data at the Buffalo River, Arkansas, as well as the value used by Green and Haggard (2001) in a previous study on the Illinois River.

The concentration data separated into baseflow and storm datasets were subjected to the three step process to evaluate changes in FACs over time (White et al., (2004) as modified by Simpson and Haggard (2016)). Log-transformed concentrations were flow-adjusted using locally weighted regression (LOESS) with an optimized smoothing parameter (ranging from 0.65 to 0.9; see Simpson and Haggard, 2016). Trends in FACs were evaluated using both a parametric and nonparametric technique. Simple linear regression was used to estimate the trend magnitude (i.e., slope) in FACs over time and determine significance of the trend with the overall F-test (α=0.05). Trends in FACs were also tested with Kendall's τ, which is insensitive to underlying distributions or extreme values in the data (Hirsch et al., 1991). However, a previous study on water quality trends found that trend significance and magnitude in FACs for both parametric and nonparametric methods (where the Sen Slope estimator was the nonparametric estimate of trend magnitude) were strongly related and generally similar in magnitude (Bailey et al., 2012).

Results and Discussion

Soluble Reactive Phosphorus Concentrations

To illustrate the separation of the water quality data, log-transformed SRP and Q for water samples from Illinois River were plotted in Figure 5. While the extreme ends of the Q range was predominantly either base or storm samples, a significant part of the range (~500-1500 cfs) have a mix of water samples collected during baseflow and storm events. This observation was consistent across all three monitoring locations, resulting from the large temporal variations in seasonal baseflow. So, a given Q might represent baseflow during the late winter and spring seasons but it might reflect storm runoff during other periods (e.g., late summer, when baseflow in the IRW is typically low). Overall, approximately 79% of the water samples were collected during baseflow conditions across these streams.

Flow-adjusted SRP concentrations revealed significant decreasing trends ($p < 0.05$; F-test and Kendall's τ) during baseflow conditions at all three sites (Figure 6). The magnitude of the decrease was greatest (12% $year^{-1}$) at Spring Creek, whereas the percent decrease was less at the two downstream sites (Osage Creek, 3% and Illinois River, 6%). Water quality data at Spring Creek was only available from 2012 to 2015 (~3.5 years), which may influence the trend interpretation. When baseflow SRP data at the other sites were constrained to the same timeframe, the decrease was similar to that observed at Spring Creek (14 and 15% $year^{-1}$ decrease ($p<0.05$) at Osage Creek and the Illinois River, respectively), which illustrates that SRP concentrations in IRW likely experienced the greatest proportional change during the last 3.5 years of our study.

When considering the major WWTPs (excluding NACA) during the period from 2009 to 2015, there is a general decreasing trend of 11% $year^{-1}$ in effluent TP concentrations. On closer inspection, Springdale's geometric mean TP was 0.32 mg L^{-1} from 2009 to 2012 and 0.25 mg L^{-1} after 2012. This decrease is incremental in comparison to that observed in an earlier study that found Springdale's effluent TP was 7.0 mg L^{-1} before October 2002 and 0.62 mg L^{-1} from October 2002 to September 2008 (Scott et al., 2011). However, Springdale's WWTP contributes much of Spring Creek's Q; the average daily effluent flow during the study was 20 cfs while average base Q in Spring Creek was 40 cfs. Since Springdale's WWTP contributes around 50% of base Q in Spring Creek, it is likely driving the SRP decreases in Spring Creek and sites further downstream.

These streams are known to be effluent-dominated, especially with respect to P concentrations during baseflow (Ekka et al., 2006; Haggard, 2010). The changes in effluent P concentrations have taken place quickly (the general trend across all WWTPs was an 11% $year^{-1}$ decrease for

2009-2015), whereas the SRP concentrations in the streams have responded somewhat slower over the same timeframe. The effluent inputs not only influence stream water P, but also P stored and available for exchange within the bottom sediments (Ekka et al., 2006; Haggard et al., 2001). The sediments can contribute to elevated SRP when the overlying water has P concentrations less than the sediment equilibrium P concentration (EPC_0, Froelich, 1988; McDowell, 2015). That is, the effluent P might decrease but the stream sediments downstream from the discharge release historic P, potentially buffering in-stream reductions of P.

Only Spring Creek exhibited a statistically significant decrease ($p<0.05$) in flow-adjusted SRP concentrations during storm events (Figure 6). Sources of P in storm flow are generally attributed to runoff from the landscape but the resuspension of P retained in stream sediments has been shown to be a potentially large P source during high flow events in the IRW (Jarvie et al., 2012). Thus, effluent P-enriched sediments that are flushed downstream during storm events may contribute to the decreasing SRP trend at Spring Creek. This effect may take longer to propagate further downstream, or the additional P sources downstream might mask the potential reductions in SRP during storm events at Osage Creek or the Illinois River.

Total Phosphorus Concentrations

Despite the improvement in baseflow SRP concentrations at all three sites, the only significant trend in flow-adjusted TP concentrations was a 9% year^{-1} decrease at Spring Creek under base conditions (Figure 7). Considering that the average SRP concentration at Spring Creek across all base samples (0.15 mg L^{-1}) makes up nearly all of the average base TP concentration (0.17 mg L^{-1}), it is unsurprising that there was a decrease in baseflow TP concentrations similar to the baseflow SRP trend. In contrast, average base SRP concentrations in Osage Creek and Illinois River (0.07 and 0.04 mg L^{-1}) are proportionally less of the average base TP concentrations (0.10

and 0.06 mg L^{-1}). Thus, SRP is a smaller proportion of TP further downstream which might explain why TP did not significantly change at Osage Creek or the Illinois River.

Interestingly, if we define the difference between TP and SRP concentrations as particulate P (PP), then PP concentrations did not change under base flow conditions at any of the sties. However, the Illinois River showed a significant increase in PP concentrations (12% $year^{-1}$) during storm flows (data not shown), whereas the other sites did not exhibit increases in PP concentrations during storm events. Reduced SRP concentrations downstream from the Springdale WWTP were likely due to abiotic (i.e., sediment) and biotic uptake which would be resuspended or scoured during storm events and transported downstream (Jarvie et al., 2012). The accumulation of effluent P in the stream bottom might promote an increase in PP concentrations during storm events downstream at the larger Illinois River, though TP concentrations do not significantly change during the study period at this site.

The lack of change in TP at the Illinois River may suggest that the effects of watershed management across IRW (e.g., poultry litter management) have not been realized yet. A lag time of years to decades in management effect for water quality is not uncommon (Meals et al., 2010). In fact, the IRW may be experiencing a lag between a P 'accumulation phase' and 'depletion phase' due to a legacy source of P (Haygarth et al., 2014; Powers et al., 2016). In addition to in-stream retention (Jarvie et al., 2012), the karst geologic features in the IRW have been shown to exhibit net P retention during storm events, where P acts like a non-conservative tracer (Jarvie et al., 2014). Furthermore, soils with historical elevated application of poultry litter provide another legacy P source (Sharpley et al., 2013). These sinks could provide a constant source of P to surface waters in the IRW which could take on the order of decades to diminish (Jarvie et al., 2014; Sharpley et al., 2009). Management of P in the IRW should take into consideration these sources in addition to the point sources evaluated in this paper and the monitoring required to detect future changes.

Conclusions

Following improvements in WWTP removal of P, three sites downstream in IRW exhibited decreasing trends in baseflow SRP from 2009 to 2015. Much of this decrease occurred after 2012, which coincides with a state regulation requiring all WWTPs in the IRW to meet a TP limit of 1 mg L^{-1} (APCEC, 2015). Unfortunately, this did not translate to a significant decrease in TP concentration at the Illinois River, despite the observed decreases of TP in the Illinois River in the 2000s (Scott et al., 2011). Legacy P in the IRW may be buffering the positive effects of watershed management, and a close examination of the potential mechanisms behind this legacy P specific to the IRW is needed in order to better focus mitigation efforts. Large strides in P management in the IRW have been made but more time to bring TP in the Illinois River down to the Oklahoma scenic river standard of 0.037 mg L^{-1} may be needed.

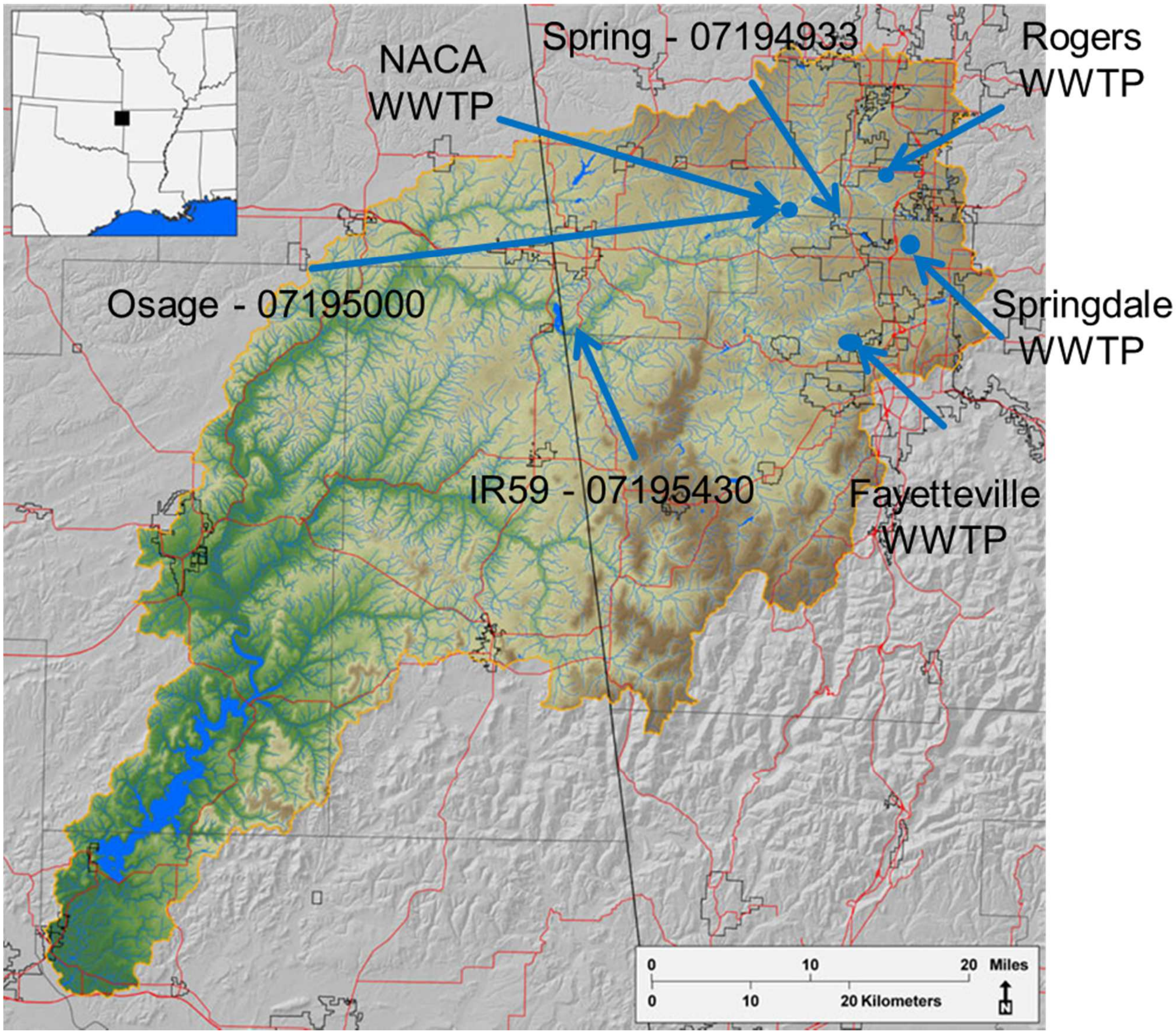

Figure 1. Map of the Illinois River Watershed with sampling sites at Spring Creek (Spring), Osage Creek (Osage) and Illinois River at Hwy 59 (IR59) and the four major wastewater treatment plants (WWTPs). USGS streamgage station numbers are given as well.

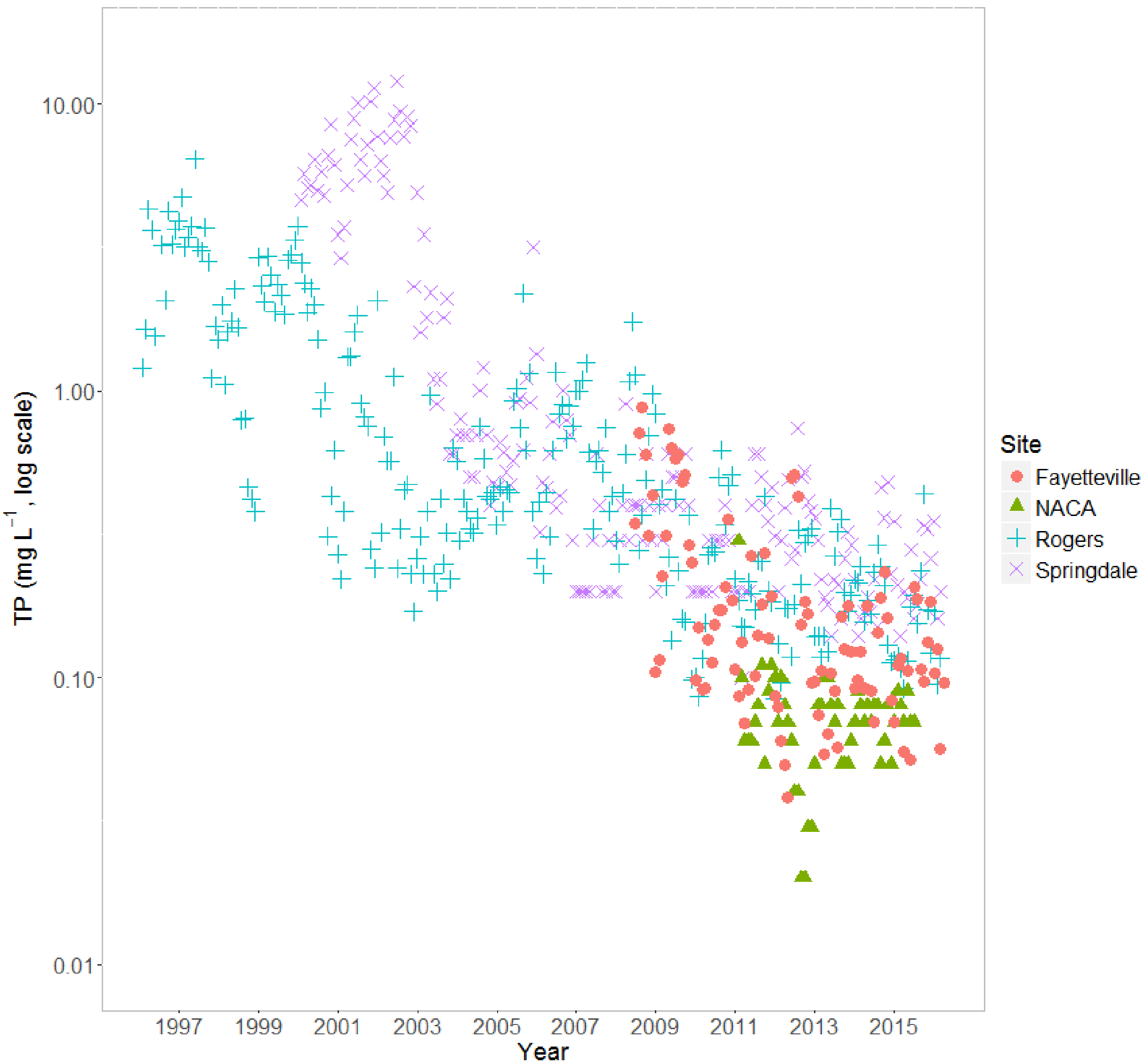

Figure 2. Monthly average total phosphorus (TP) concentrations in the effluent at four major WWTPs in the IRW; period of record is variable between plants, and data was graciously provided by personnel from each treatment facility.

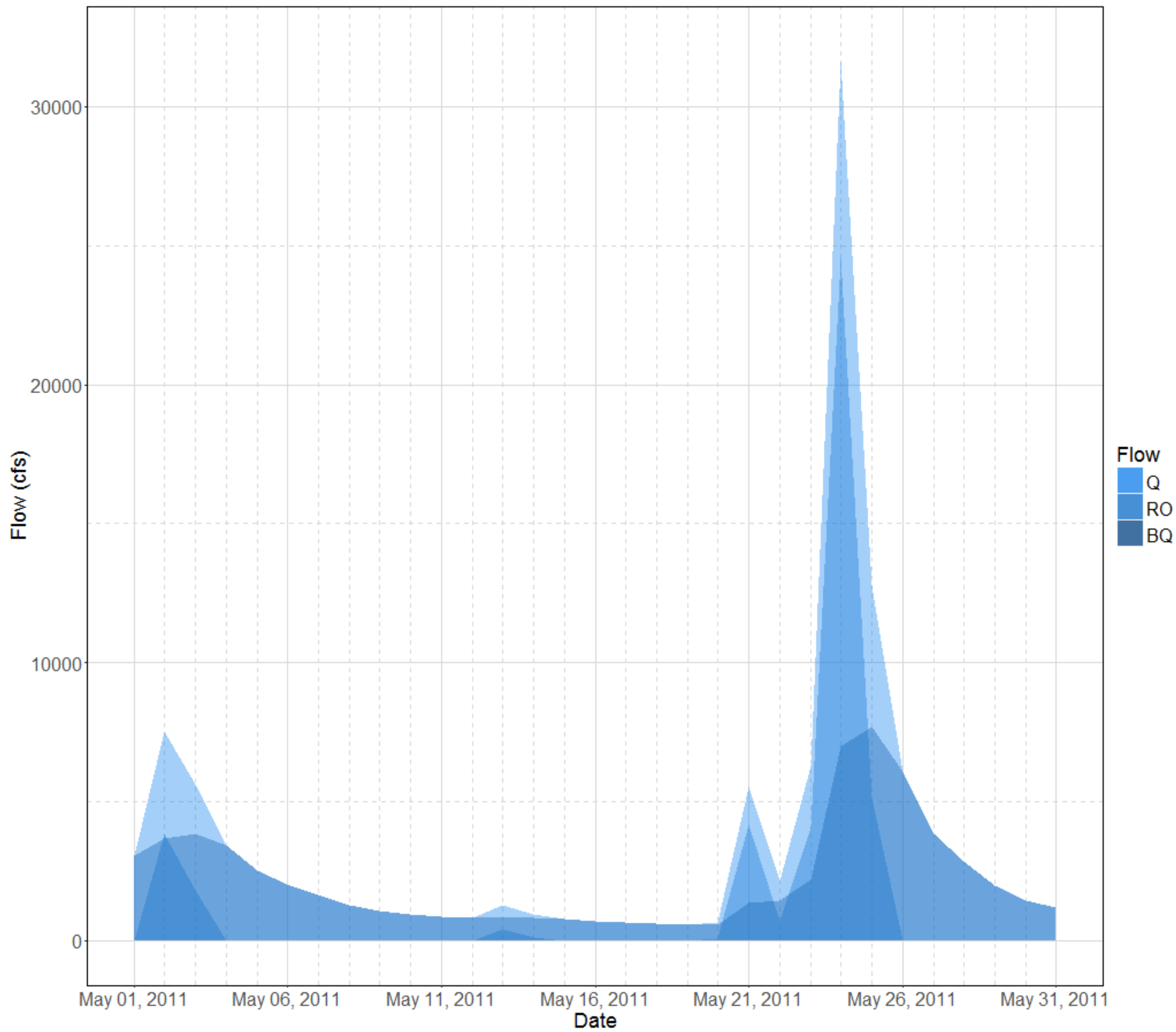

Figure 3. Hydrograph separation for the month of May 2011 at the Illinois River using the Eckhardt baseflow filter; the mean daily streamflow (Q; cfs) is separated into two components: baseflow (BQ) and a runoff component (RO).

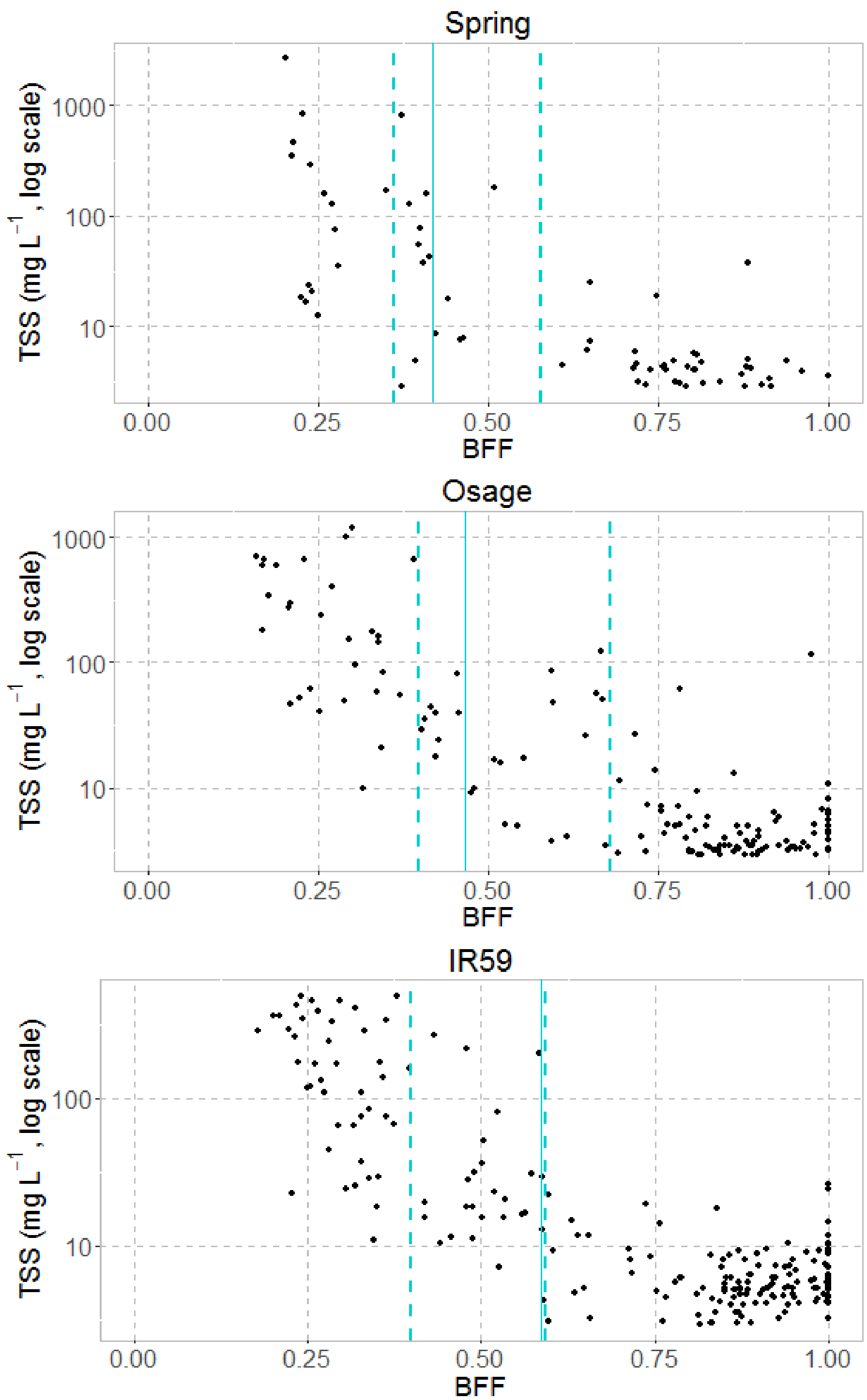

Figure 4. Changepoints in baseflow fraction (BFF) identified with TSS data using nonparametric changepoint analysis (NCPA; see text) at Spring Creek, Osage Creek, and Illinois River at Hwy 59 (IR59). The solid line is the identified changepoint while the dashed lines are the 95% confidence interval estimated via bootstrapping. Data points with BFF lower than 0.6 were termed as storm data while points greater than 0.6 were termed as baseflow data.

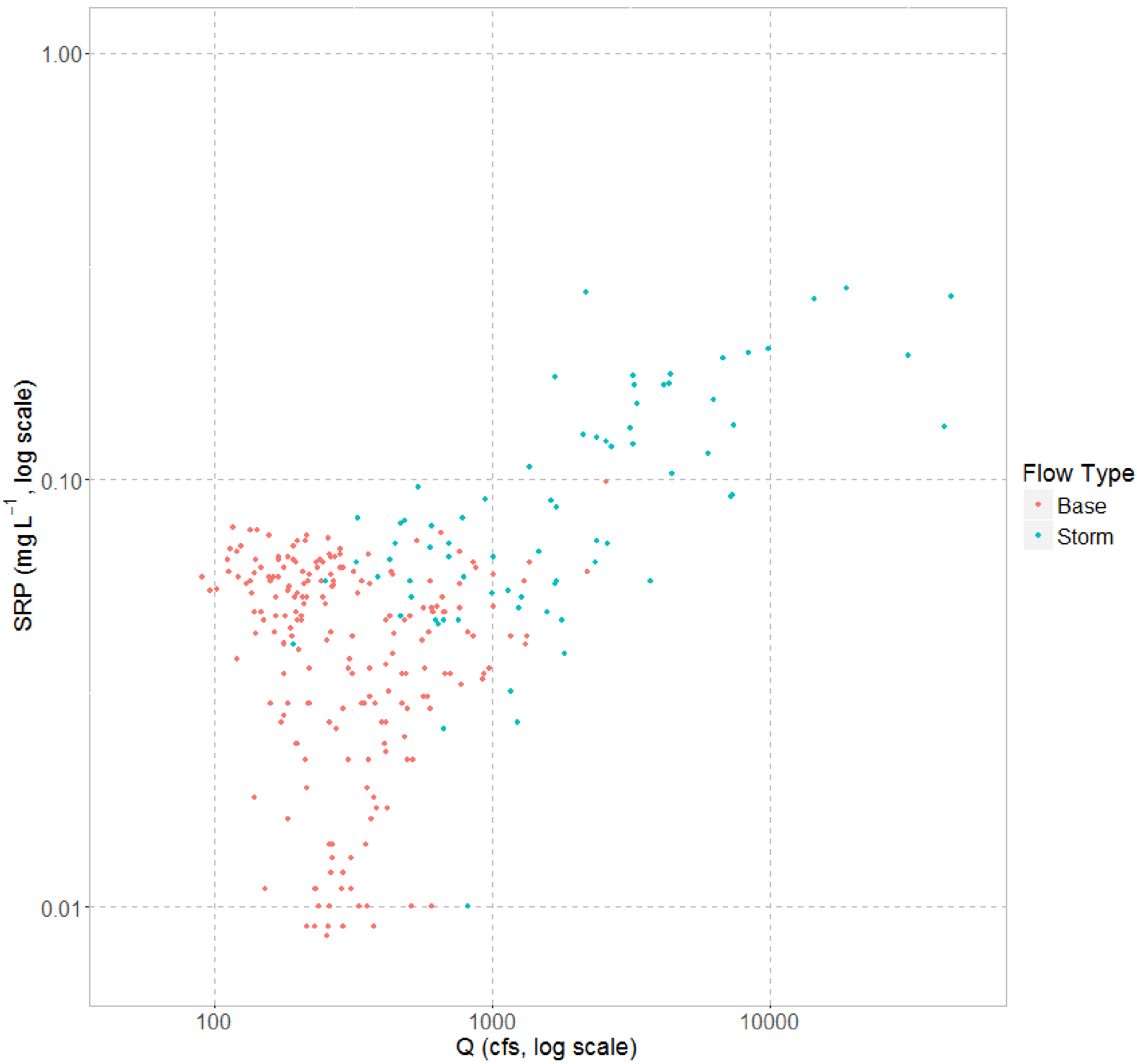

Figure 5. Separation of soluble reactive phosphorus (SRP) concentration data at the Illinois River into base and storm samples based on the BFF changepoint described in text.

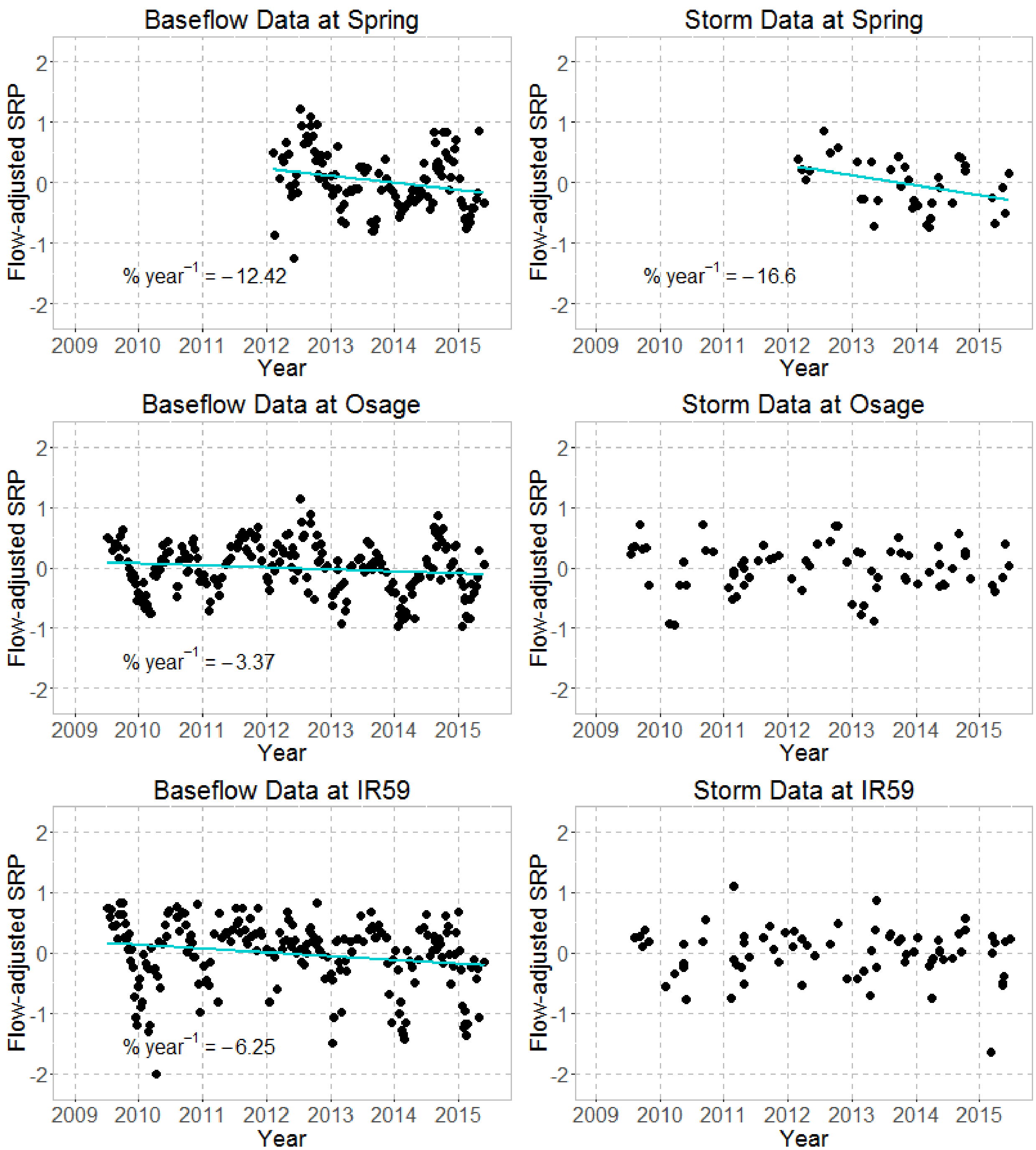

Figure 6. Trend analyses of soluble reactive phosphorus (SRP) data under both baseflow and storm conditions at Spring Creek (Spring), Osage Creek (Osage), and Illinois River at Hwy 59 (IR59). Significant trends (F-test, $p<0.05$) are shown with a linear regression and trend magnitude is reported as percent change in concentration per year.

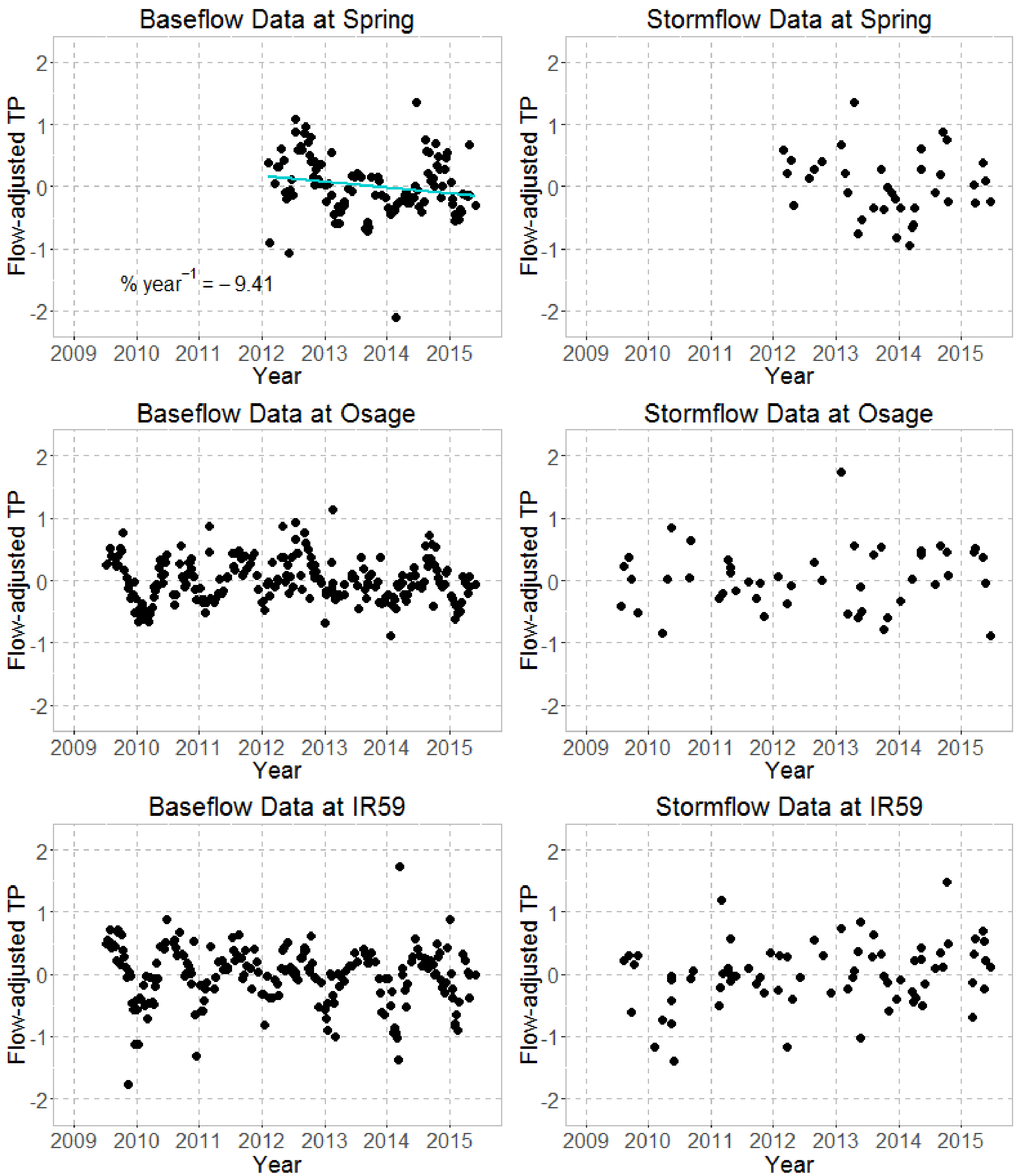

Figure 7. Trend analyses of TP data under both baseflow and storm conditions at Spring Creek (Spring), Osage Creek (Osage), and Illinois River at Hwy 59 (IR59). Significant trends (F-test, $p<0.05$) are shown with a linear regression and trend magnitude is reported as percent change in concentration per year.

Works Cited

Arkansas Pollution Control and Ecology Commission (APCEC). 2015. Regulation No. 2: Regulation establishing water quality standards for surface waters of the state of Arkansas. Accessed at https://www.adeq.state.ar.us/regs/files/reg02_final_151114.pdf

Arkansas Pollution Control and Ecology Commission (APCEC). 2015. Regulation No. 6: Regulations for state administration of the National Pollutant Discharge Elimination System (NPDES). Accessed at https://www.adeq.state.ar.us/regs/files/reg06_final_150918.pdf

AWRC. 2016. AWRC Water Quality Laboratory. Accessed at http://arkansas-water-center.uark.edu/waterqualitylab.php on April 21, 2016.

Bailey, B.W., B.E. Haggard, and L.B. Massey. 2012. Water quality trends across select 319 monitoring sites in Northwest Arkansas. Arkansas Water Resources Center, Fayetteville, Arkansas, MSC Publication 365: 50 pp.

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. Ecological Applications 8(3): 559-568.

CAST. 2006. Arkansas Watershed Information System: a module of the Arkansas Automated Reporting and Mapping System. Accessed at: http://watersheds.cast.uark.edu/index.php

Chapman, T.G. 1991. Comment on “Evaluation of Automated Techniques for Base Flow and Recession Analyses” by R.J. Nathan and T.A. McMahon. Water Resources Research 27(7): 1783-1784.

Dodds, W.K., W.W. Bouska, J.L. Eitzmann, T.J. Pilger, K.L. Pitts, A.J. Riley, J.T. Schloesser, and D.J. Thornbrugh. 2009. Eutrophication of U.S. freshwaters: Analysis of potential economic damages. Environmental Science and Technology 43(1): 12-19.

Eckhardt, K. 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes 19: 507-515.

Eckhardt, K. 2008. A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. Journal of Hydrology 352: 168-173.

Ekka, S.A., B.E. Haggard, M.D. Matlock, I. Chaubey. 2006. Dissolved phosphorus concentrations and sediment interactions in effluent-dominated Ozark streams. Ecological Engineering 26: 375-391.

Froelich, P.N. 1988. Kinetic control of dissolved phosphate in natural rivers and estuaries: a primer on the phosphate buffer mechanism. Limnology and Oceanography 33(4): 649 - 668.

Green, W.R., and B.E. Haggard. 2001. Phosphorus and nitrogen concentrations and loads at Illinois River South of Siloam Springs, Arkansas, 1997-1999. U.S. Geological Survey. Water-Resources Investigations Report 01-4217.

Haggard, B.E., D.E. Storm, R.D. Tejral, Y.A. Popova, V.G. Keyworth, E.H. Stanley. 2001. Stream nutrient retention in three northeastern Oklahoma agricultural catchments. Transactions of the ASAE 44(3): 597-605.

Haggard, B.E. 2010. Phosphorus concentrations, loads, and sources within the Illinois River drainage area, northwest Arkansas, 1997-2008. Journal of Environmental Quality 39: 2113-2120.

Haygarth, P.M., et al. 2014. Sustainable phosphorus management and the need for a long-term perspective: the legacy hypothesis. Environmental Science and Technology 48: 8417-8419.

Herron, S., A.N. Sharpley, S. Watkins, and M. Daniels. 2012. Poultry litter management in the Illinois River Watershed of Arkansas and Oklahoma. Cooperative Extension Service, Division of Agriculture, University of Arkansas. Fact Sheet FSA 9535, 4 p. Accessed at http://www.uaex.edu/publications/pdf/FSA-9535.pdf

Hirsch, R.M., R.B. Alexander, and R.A. Smith. 1991. Selection of methods for the detection and estimation of trends in water quality. Water Resources Research 27(5): 803-813.

Jarvie, H.P., A.N. Sharpley, J.T. Scott, B.E. Haggard, M.J. Bowes, and L.B. Massey. 2012. Within-river phosphorus retention: accounting for a missing piece in the watershed phosphorus puzzle. Environmental Science and Technology 46: 13284-13292.

Jarvie, H.P., A.N. Sharpley, V. Brahana, T. Simmons, A. Price, C. Neal, A.J. Lawlor, D. Sleep, S. Thacker, and B.E. Haggard. 2014. Phosphorus retention and remobilization along hydrological pathways in karst terrain. Environmental Science and Technology 48: 4860-4868.

McDowell, R.W. 2015. Relationships between sediment chemistry, equilibrium phosphorus concentrations, and phosphorus concentrations at baseflow in rivers of the New Zealand River Water Quality Network. Journal of Environmental Quality 44: 921-929.

Meals, D.W., S.A. Dressing, T.E. Davenport. 2010. Lag time in water quality response to best management practices: a review. Journal of Environmental Quality 39: 85-96.

Nathan, R.J. and T.A. McMahon. 1990. Evaluation of automated techniques for base flow and recession analyses. Water Resources Research 26(7): 1465-1473.

OCC. 2010. Illinois River Watershed Based Plan. Oklahoma Conservation Commission (OCC) Water Quality Division. Accessed at: http://www.ok.gov/conservation/documents/Illinois%20River%20Watershed-based%20plan%20approved%202_11_%202011.pdf

OWRB. 2002. Water Quality Standards. Oklahoma Water Resources Board, Oklahoma City, OK, USA.

Powers, S.M., et al. 2016. Long-term accumulation and transport of anthropogenic phosphorus in three river basins. Nature Geoscience. doi: 10.1038/NGEO2693

Qian, S.S., R.S. King, and C.J. Richardson. 2003. Two statistical methods for the detection of environmental thresholds. Ecological Modelling 166: 87-97.

Scott, E.E., Z.P. Simpson, and B.E. Haggard. 2015. Constituent loads and trends in the Upper Illinois River Watershed and Upper White River Basin. Arkansas Water Resources Center, Fayetteville, Arkansas, MSC Publication 377: 89 pp.

Scott, J.T., B.E. Haggard, A.N. Sharpley, and J.J. Romeis. 2011. Change point analysis of phosphorus trends in the Illinois River (Oklahoma) demonstrates the effects of watershed management. Journal of Environmental Quality 40: 1249 - 1256.

Sharpley, A.N., J.L. Weld, D.B. Beegle, P.J.A. Kleinman, W.J. Gburek, P.A. Moore Jr., and G. Mullins. 2003. Development of phosphorus indices for nutrient management planning strategies in the United States. Journal of Soil and Water Conservation 58(3): 137 - 152.

Sharpley, A.N., P.J.A. Kleinman, P. Jordan, L. Bergström, and A.L. Allen. 2009. Evaluating the success of phosphorus management from field to watershed. Journal of Environmental Quality 38: 1981-1988.

Sharpley, A.N., H.P. Jarvie, A. Buda, L. May, B. Spears, and P. Kleinman. 2013. Phosphorus legacy: overcoming the effects of past management practices to mitigate future water quality impairment. Journal of Environmental Quality. 42: 1308-1326.

Simpson, Z.P. and B.E. Haggard. 2016. An optimized procedure for flow-adjustment of constituent concentrations for trend analysis. In preparation.

Vadas, P.A., L.W. Good, P.A. Moore Jr., and N. Widman. 2009. Estimating phosphorus loss in runoff from manure and fertilizer for a phosphorus loss quantification tool. Journal of Environmental Quality 38: 1645 - 1653.

White, K.L., B.E. Haggard, and I. Chaubey. 2004. Water quality at the Buffalo National River, Arkansas, 1991-2001. Transactions of the American Society of Agricultural Engineers 47(2): 407-417.

III. Relating Water Quality in Beaver Lake to Nutrient Trends in Watershed Inflows

Abstract

Beaver Lake (northwest Arkansas) has recently adopted an effects-based water quality standard for algal productivity where growing season geometric mean chlorophyll-a (chl-a) must not exceed 8 μg L^{-1} and annual average Secchi transparency must not fall below 1.1 m. However, Beaver Lake is likely to violate this standard since chl-a concentrations have been increasing near the assessment location. We evaluated watershed nutrient inputs to Beaver Lake from 2009 to 2015 via trend analysis of monitoring data collected at the three main inflows, as well as examine chl-a and nutrient data collected at two sites in Beaver Lake. Flow-adjusted concentrations of soluble reactive phosphorus (SRP) and total P (TP) have increased at Richland Creek (5.2 and 7.4% $year^{-1}$, respectively) while a 3.7% $year^{-1}$ decrease has occurred at War Eagle Creek. Total nitrogen (TN) concentrations have increased at Richland Creek and White River (7.3 and 3.8% $year^{-1}$, respectively); however, nitrogen (N) concentrations (both NO_3-N and TN) exhibit a nonlinear pattern across all three sites, where N increased from 2009 through 2012 and have potentially decreased or maintained from 2013 through 2015. The pattern in N is likely tied to hydroclimatic factors, such as an extensive drought that occurred in 2012. Data collected at Beaver Lake illustrated an increasing trend in both growing season geometric mean chl-a (0.24 μg L^{-1} $year^{-1}$) and annual average TN (0.02 mg L^{-1} $year^{-1}$) from 2001 to 2015 at the site closest to a drinking water supply intake (Lowell, Arkansas). This study provides important insight to nutrient concentrations in the Beaver Lake watershed for the past 6 years (2009 through 2015), but longer records of data are needed to fully understand short-term fluctuations and long-term persistent trends of N and P in this basin. Watershed-scale nutrient management and

hydroclimatic factors are key considerations for understanding variations in algal productivity in Beaver Lake.

Introduction

The eutrophication of freshwater bodies is a pervasive problem throughout the U.S. (Brown and Froemke, 2012), particularly in agricultural watersheds where nutrient inputs are elevated. A study of the economic impact eutrophication has on U.S. freshwaters estimates an annual loss of \$2.2 billion when accounting for impacts such as loss of aesthetic value, loss of biodiversity, and water treatment (Dodds et al., 2009). The eutrophication of drinking water supply reservoirs presents concerns over increased treatment costs, taste and odor issues, disinfection by-products (DBPs) and even harmful algal blooms, requiring water treatment facilities to use non-conventional treatment to provide drinking water (Walker Jr., 1983). In northwest Arkansas, the Upper White River Basin (UWRB) is unique in that it also contains northwest Arkansas's primary drinking water source, Beaver Lake, and Beaver Water District has recently shifted to more non-conventional pretreatment processes (i.e., chlorine-dioxide pre-disinfection) to address eutrophication and potential DBP formation (Beaver Water District, 2014).

Recently, an effects-based standard for eutrophication in Beaver Lake was adopted, where reservoir eutrophication is regulated via algal biomass (i.e., chlorophyll-a concentrations) and clarity. The standard states that the growing season (i.e., May through October) geometric mean chlorophyll-a (chl-a) concentration in Beaver Lake near Hickory Creek shall not exceed 8 $\mu g\ L^{-1}$ and that the annual average Secchi transparency shall not be less than 1.1 m (APCEC, 2015). However, Scott and Haggard (2015) point out the numerous issues with this standard; for instance, the assessment location is in the riverine-transition zone and chl-a concentrations have generally been increasing in this zone at Beaver Lake. Reservoirs that are at risk of eutrophication typically exhibit elevated algal biomass in the riverine and transition zones (e.g., Scott et al., 2009; Cooke et al., 2011).

The proposed assessment methodology is that the geometric mean chl-a concentration at Hickory Creek cannot exceed the state's water quality standard in more than two out of five years (APCEC, 2016). According to collected and modeled chl-a data from 2001 to 2014, there is a 60% probability that a five-year assessment at Hickory Creek would result in violation (Scott and Haggard, 2015). Thus, it seems likely that Beaver Lake will be listed as impaired after the standard has been implemented and assessment occurs by the Arkansas Department of Environmental Quality (ADEQ), which is likely an unintended consequence. However, it is important to understand the reasons why chl-a concentrations have increased in Beaver Lake, including external processes that might be driving the high productivity.

The UWRB is mostly forested but is still influenced by pasture and urban land use, where greater percent pasture and urban land use within drainage basins have been linked to greater in-stream nitrogen (N) and phosphorus (P) concentrations (Haggard et al., 2003; Migliaccio et al., 2007; Giovannetti et al., 2013). Watershed nutrient inputs delivered to drinking water supply reservoirs and internal processes (i.e., P release from bottom sediments) influence algal productivity in the water column. The potential future 303(d) listing of Beaver Lake for its water quality standard violation will essentially open the opportunity to start developing the total maximum daily load (TMDL), focusing on watershed nutrient inputs because internal loading is relatively low (Sen et al., 2007).

In the present study, we examine trends in N, P, and chl-a concentrations at two sites of interest at Beaver Lake, as well as N and P trends in the three main tributaries to Beaver Lake using monitoring data from 2009 to 2015. We wanted to know if changes in the N, P, and chl-a within the transitional zone of Beaver Lake were linked to changes in watershed inputs, suggesting that increased nutrient transport into Beaver Lake resulted in increased productivity and potential 303(d) listing of the water body. We hypothesized that patterns in algal biomass at Beaver Lake were closely tied to variability in nutrient availability from stream inputs.

Methods

Site Description

The UWRB (HUC 11010001) lies in northwest Arkansas with a small portion in Missouri and abuts the Illinois River Watershed to the east within the major metropolitan areas of Fayetteville, Springdale, and Rogers, Arkansas. The watershed drains approximately 574,718 ha before flowing into Missouri. In 2006, land use was mostly forest (64%) and pasture (23%) with some urban (4%) (CAST, 2006). Madison and Carroll counties, which are nearly completely within the watershed, have seen little population growth from 2000 to 2010; however, the major metropolitan areas within Benton and Washington counties, have seen an increase in population of approximately 44% and 29% from 2000 to 2010, respectively (2010 Census Data). The White River is impounded to form Beaver Lake (approximately 11,396 ha) which serves as the primary water supply for northwest Arkansas. On average, UWRB gets 1230 mm year^{-1} of precipitation; however, in 2012, there was only 797 mm of precipitation (NADP, 2016), resulting in extreme drought throughout the growing season of 2012 (via Palmer Drought Severity Index; NOAA, 2016).

The three main tributaries to Beaver Lake are the White River, Richland Creek, and War Eagle Creek. The West Fork of the White River contributes roughly 30% of the annual Q to the White

River while the remainder comes from the outflow of Lake Sequoyah, an impoundment of the Middle Fork and eastern mainstem of the White River. Additionally, one of Fayetteville's wastewater treatment plants (WWTP; Paul Noland Wastewater Treatment Facility) is located on the White River, just downstream from the sampling site on the White River. The plant has an average discharge of 48,000 m^3 day^{-1} and has historically operated under a monthly average total phosphorus (TP) standard of 1 mg L^{-1}; however, effluent concentrations were consistently under 0.5 mg L^{-1} during our study. Another WWTP, although smaller, is located in Huntsville and discharges upstream from the sampling site on War Eagle Creek.

Water samples were collected at USGS gaging stations at West Fork of the White River (West Fork; USGS 07048550), White River near Fayetteville, AR (White River; USGS 07048600), Richland Creek at Hwy 45 (Richland Creek; USGS 07048800), and War Eagle Creek (USGS 07049000), shown in Figure 1. Grab samples were collected from the stream's centroid of flow on a near weekly basis including both baseflow and storm events from bridge access using an alpha style horizontal sampler. Water samples were analyzed at the Arkansas Water Resources Center (AWRC) certified water quality lab for nitrate-nitrogen (NO_3), total nitrogen (TN), soluble reactive phosphorus (SRP), and TP using approved standard methods for the analysis of water samples (AWRC, 2016). Data at each stream was collected from July 2009 to June 2015 (6 years) except at Richland Creek, where data was only available through April 2015 due to backwater conditions in Beaver Lake. The stream gage at Richland Creek was subsequently moved upstream at the end of this study. Discharge data (cfs) were available from the USGS National Water Information System (NWIS) for each site during this study period.

In addition to water samples from the inflows collected by AWRC, water samples from Beaver Lake were collected by the USGS at five sites from the riverine zone down the reservoir to the dam. Grab samples were collected about 2 m below surface and analyzed for chl-a, TN, and TP at the USGS National Water Quality Lab (Denver, Colorado). Our analysis of the lake data focused on two sites, including Hickory Creek and Lowell. The site at Hickory Creek on Beaver Lake is where the geometric mean chl-a standard (8 $\mu g\ L^{-1}$) during the growing season (May through October) applies; the assessment proposed by ADEQ is a violation will result if exceeded more than two times within a five year period (ADEQ, 2016). The site at Lowell on Beaver Lake is in close proximity to the intake of Beaver Water District, which provides drinking water to approximately 400,000 residents as well as multiple industries. Nutrient and chl-a data were available from the USGS NWIS since 2001 at Lowell and since 2009 at Hickory Creek.

Analysis

Daily mean discharge (Q) was used to generate the baseflow discharge (BQ) record via hydrograph separation (Eckhardt, 2005). The procedure, originally developed by Nathan and McMahon (1990), features an improvement on the BQ filter algorithm (Chapman, 1991), and produces hydrologically sound results (Eckhardt, 2008). The filter values recommended for perennial streams with porous aquifers were used (Nathan and McMahon, 1990; Eckhardt, 2005), since the UWRB is characterized by limestone geology with some karst features. Unfortunately, one site in our study (White River) had too many gaps in the later part of the hydrological record to allow for analysis of BQ. Therefore, these water samples at White River were separated into baseflow and storm event samples evaluating field sheets and sampling conditions at other sampling sites in the UWRB monitoring program (e.g., West Fork of the White River).

A baseflow fraction (BFF; BQ/Q) was determined for each sample collected at sites where BQ data was available. Based on change point analysis of water quality data collected in the adjacent Illinois River Watershed, a BFF value of 0.6 was used to separate data into baseflow data and storm data (Simpson and Haggard, 2016A). That is, data with BFF ≥ 0.6 (i.e., 60% or more of Q is BQ) were classified as baseflow data, while data with BFF < 0.6 were considered as storm data. Approximately 70% of water samples collected in this study were during baseflow conditions.

All water quality data, as well as data split into baseflow and storm data, were subjected to the three step process to evaluate changes in flow-adjusted concentrations (FACs) over time (White et al., (2004) as modified by Simpson and Haggard (2016B)). Log-transformed concentrations were flow-adjusted using locally weighted regression (LOESS) with an optimized smoothing parameter, which ranged from 0.3 to 1.0 (see Simpson and Haggard, 2016B). Trends in FACs were evaluated using both a parametric and nonparametric technique (Helsel and Hirsch, 2002). Simple linear regression was used to estimate trend magnitude (i.e., slope) in FACs over time and determine significance of the trend with the overall F-test (α=0.05). Trends were also evaluated with Kendall's τ, a rank-correlation statistic that is insensitive to underlying distributions or extreme values in the data (Hirsch et al., 1991). A previous study on water quality trends found that trend interpretations using both methods generally found similar trend magnitudes and degree of significance (Bailey et al., 2012).

It should be noted here that a key assumption we make in the trend analysis of this data is that trends are only monotonic (i.e., only increasing or only decreasing over the entire period). However, water quality (i.e., FACs) might show subtle changes over time in response to variations in climate (e.g., drought) or other factors during the study period. Therefore, LOESS was used to evaluate subtle shifts in FACs over time that might not be expressed monotonically.

For the analysis of Beaver Lake data, the growing season geometric means of chl-a were evaluated for changes in time. We analyzed chl-a at Hickory Creek, since this is the site of the 8 $\mu g\ L^{-1}$ standard assessment, and we analyzed data at Lowell to gain perspective on chl-a concentrations on a longer time scale. Additionally, nutrient data at Lowell were evaluated on annual average basis. Simple linear regression was used to determine the change in the central tendency of these parameters over time.

Results and Discussion

UWRB Inputs: Soluble Reactive Phosphorus

Baseflow concentrations of SRP were generally low for this study period with geometric mean SRP concentrations of 0.004, 0.006, and 0.005 mg L^{-1} in Richland Creek, War Eagle Creek and White River, respectively. These streams reflect the lower range of baseflow SRP seen across multiple sites in the UWRB (0.003 – 0.021 mg L^{-1}; Giovannetti et al., 2013). SRP concentrations in storm events at the tributaries were considerably higher (0.013 – 0.019 mg L^{-1}) but similar to the range (0.010 – 0.054 mg L^{-1}) reported by Giovannetti et al. (2013). It is well known that stream nutrients can correlate with the percent pasture in the drainage area, especially in the UWRB (Haggard et al., 2003; Migliaccio et al., 2007; Giovannetti et al., 2013). However, the drainage areas of these streams are largely forested (>50% forest in each basin) and represent a larger subwatershed area as the monitoring sites are located near the entrance to Beaver Lake. Indeed, annual nutrient export (i.e., mass exported per subwatershed area) is significantly greater in the smaller catchments in UWRB (Haggard et al., 2003). Larger drainage areas may buffer the critical source areas that are contributing much of the nutrient loss (McDowell et al., 2004).

No trends were detected in the SRP concentrations of water samples collected during only baseflow conditions; however, War Eagle Creek showed an 11.7% $year^{-1}$ decrease in storm SRP (Table 1). When analyzing all data, this same trend was dampened to a 3.8% $year^{-1}$ decrease at War Eagle Creek, illustrating that decreases in SRP at this site are largely due to decreases in SRP during elevated flows (Figure 2). Interestingly, Richland Creek only showed a significant increasing trend when analyzing all data (5.4% $year^{-1}$), while no trend was detected in either baseflow or storm data. This may be an artifact of greater statistical power in trend analysis when using more observations.

UWRB Inputs: Total Phosphorus

TP concentrations under baseflow conditions were variable across all three sites where the geometric means ranged from (0.017 – 0.031 mg L^{-1}), similar to the range reported for sites in the UWRB not influenced by effluent discharge (0.009 – 0.070 mg L^{-1}; Giovannetti et al., 2013). Though only one site in our study was impacted by a major WWTP (War Eagle Creek), point sources like WWTPs have a large impact on nutrient concentrations under baseflow conditions, where the near-constant source can elevate concentrations within the stream while likely playing a smaller role in the annual TP load (Stamm et al., 2013). Additionally, it is likely that excess nutrients from the Fayetteville WWTP effluent discharged into the White River are not being sequestered before reaching Beaver Lake (Hufhines et al., 2011). However, the major WWTP in the UWRB (Fayetteville) has historically managed effluent TP concentrations below 0.5 mg L^{-1} (e.g., was generally less than 0.4 mg L^{-1} from 2006 to 2007; Hufhines et al., 2011) as well as during the current study period (Figure 3), which is less than the effluent TP limit of 1.00 mg L^{-1}.

We did not collect discharge data from the Huntsville WWTP, however, our site on War Eagle Creek, downstream from the WWTP, maintained relatively low TP concentrations in baseflow (geometric mean of 0.023 mg L^{-1}) for the period.

The only site to exhibit significant trends in TP for the study was Richland Creek (Figure 2). Baseflow TP increased 9.1% $year^{-1}$ which resulted in a 7.6% $year^{-1}$ increase when considering all data (Table 1). Most of the watershed has not had significant changes in TP, which could be due to initiatives like the P index (DeLaune et al., 2004) stabilizing excess P inputs on the landscape. While there was a statistically significant trend in TP at Richland Creek, the geometric mean TP concentration across all water samples was relatively low (0.027 mg L^{-1}); proportional increases in TP at this site are less likely to significantly alter nutrient dynamics within Beaver Lake. It is unclear what sources may drive increases in TP at Richland Creek. Potentially, groundwater stores of P as well as historical inputs of watershed P (e.g., land-applied poultry litter) may provide a legacy P source (Jarvie et al., 2014).

UWRB Inputs: Nitrate

Most of the TN fraction in UWRB was made up by NO_3, which varied across sites (58 to 87%, on average). Richland Creek and the White River had geometric mean NO_3 concentrations of 0.42 and 0.28 mg L^{-1}, respectively, under baseflow conditions, reflecting the lower part of the range reported for baseflow NO_3 in the UWRB (0.05 – 2.28 mg L^{-1}; Giovannetti et al., 2013). However, War Eagle Creek had a geometric mean NO_3 concentration of 1.40 mg L^{-1} for the study period. This elevated concentration is similar to that reported elsewhere for streams in the War Eagle Creek watershed (Migliaccio et al., 2007; Giovannetti et al., 2013). Subwatersheds within UWRB exhibit a positive correlation between stream NO_3 concentrations and percent pasture (Giovannetti et al., 2013), which may explain the elevated concentrations in War Eagle Creek, which is approximately 35% pasture.

Monotonic trends for NO_3 indicated that there was a significant increase at Richland Creek under baseflow conditions (11.3% year^{-1}; Table 1). Additionally, a statistically significant decrease in storm NO_3 concentrations (6.2% year^{-1}) occurred at War Eagle Creek. When analyzing all data at these sites, the increasing trend at Richland Creek was still evident (8.4% year^{-1}; Figure 4) but no trend in NO_3 concentrations was apparent at War Eagle Creek despite the decrease in NO_3 concentrations during storm events. The White River showed no significant changes during this study period. These analyses assume monotonic changes in the constituent, however the relationship between flow-adjusted NO_3 concentrations and time at each site illustrated a nonlinear pattern (Figure 4). Following the smoother regression to the data, there is an apparent increase from the early part of the period to late 2012/2013, accompanied by an increase in the variance of the data. After this period, NO_3 at War Eagle Creek and White River likely decreased again; meanwhile, this shift may not be evident at Richland Creek.

The 2012 period in this study experienced extreme drought, as evidenced by the large dip in streamflow at each site (Figure 5). Hydroclimatic variability may affect the transport of NO_3 through groundwater sources. NO_3 can be highly mobile in sufficiently wet soils due to lack of binding forces, which results in leaching. It is considered that this mechanism may explain the elevated baseflow NO_3 concentrations in subwatersheds with extensive pasture land use, especially in the War Eagle Creek watershed (Migliaccio et al., 2007); the karst geology in the UWRB could supply more of the stream flow from diffuse sources within the groundwater that contain elevated concentrations of NO_3 (Boyer and Pasquarrell, 1995; Green and Haggard, 2001, Bowes et al., 2015). Another effect of dry periods, especially during severe drought conditions like in 2012, is that excess organic N could build up on soil surfaces when little rain-driven percolation or runoff has occurred, leading to a large NO_3 source delivered when sufficient rain relieves the dry conditions (Boyer and Pasquarrell, 1995; Mosley, 2015). It is possible that the UWRB experienced a buildup of N during the 2012 drought, which could explain the relatively low flow-adjusted NO_3 concentrations during this period (Figure 4), and flushed out excess N immediately after the drought (see Mosley, 2015 and references therein).

UWRB Inputs: Total Nitrogen

Baseflow TN concentrations in the UWRB follow the same pattern as NO_3, with geometric means ranging from 0.50 mg L^{-1} to 1.54 mg L^{-1}, similar to the range reported for sites across UWRB (0.17 – 2.33 mg L^{-1}; Giovannetti et al., 2013). The highest TN concentrations also occurred at War Eagle Creek (1.54 mg L^{-1} geometric mean), consistent with the relatively high TN concentrations recorded in this watershed previously (1 to 4 mg L^{-1} in sites closer to outlet of watershed; Migliaccio et al., 2007).

Two sites reflected a monotonic change in baseflow TN concentrations (Table 1). Baseflow TN increased 8.3% $year^{-1}$ at Richland Creek and 6.3% $year^{-1}$ at White River. Additionally, storm event TN increased 5.2% $year^{-1}$ at Richland Creek. When considering all data, TN increased at both Richland Creek and White River by 7.3 and 3.8% $year^{-1}$, respectively (Figure 4). There was no statistically significant change in TN at War Eagle Creek. While increasing trends are evident for both NO_3 and TN at Richland Creek, the discrepancy at White River where only TN showed changes could be attributed to additional inputs of N in organic forms. Similar to NO_3, TN likely underwent nonlinear changes during the study period (Figure 4). Again, flow-adjusted TN concentrations peaked during the late 2012/2013 period, after a major drought had occurred, likely reflecting groundwater influence as well as concentrated nutrient cycling under the low flows. Droughts are important hydrological events that can greatly shift watershed nutrient concentrations (Mosely, 2015), and need to be considered when examining algal productivity within reservoirs such as Beaver Lake.

Beaver Lake

Water samples collected at Hickory Creek had chl-a concentrations ranging from 1.8 to 17.5 µg L^{-1} from 2009 to 2015, whereas geometric mean chl-a concentrations were from 6.3 to 12.3 µg L^{-1} during the growing season (May through October; Figure 6). Five of these seven years exceeded the 8 µg L^{-1} standard for chl-a at Beaver Lake, where the assessment location is at Hickory Creek. Hickory Creek lies in the riverine-transition zone of Beaver Lake downstream from the confluence of all major tributaries. The riverine-transition zone of reservoirs with elevated nutrient inputs, such as at Beaver Lake, are highly productive areas (Cooke et al., 2011; Scott et al., 2009) and are often sites of intense biogeochemical processes (e.g., nitrogen-fixation; Scott et al., 2008). Algal productivity in this zone is fueled by the nutrient inputs from the three tributaries and declines (i.e., along a trophic gradient) further down the reservoir towards the outlet (Scott and Haggard, 2015).

At the Lowell site, geometric mean chl-a concentrations ranged from 1.4 to 9.5 µg L^{-1} and increased at a rate of 0.24 µg L^{-1} $year^{-1}$ from 2001 to 2015 (Figure 7). Accompanying this trend, annual average TN increased at a rate 0.02 mg L^{-1} $year^{-1}$, where concentrations ranged from 0.59 to 1.00 mg L^{-1}. Annual average TP concentrations ranged from 0.016 to 0.037 mg L^{-1}, typical of mesotrophic to eutrophic lakes (Nürnberg, 1996) and similar to another lake impacted by agricultural P in eastern Oklahoma (Lake Tenkiller; Cooke et al., 2011). However, no trend was apparent in TP at Lowell. Increased productivity in Beaver Lake may drive such high N removal in the water column that the ecosystem has become sensitive to additional N loadings (Scott and

McCarthy, 2010; Finlay et al., 2013), which may be supported by the coincidental increases in TN and chl-a at Lowell. More and more evidence points to N and P co-limitation (Sterner, 2008), meaning that management should consider P as well as N inputs to water bodies.

Attention should also be placed on other hydrological and climatic variables when considering lake water quality data. Water residence time and temperature can provoke greater algal productivity (Paerl and Huisman, 2008), which may explain peaks in chl-a concentrations, such as in 2012 (Figure 6 and 7). In fact, droughts in Beaver Lake's drainage area have been tied to elevated 2-methylisoborneol (MIB) concentrations, a taste-and-odor compound produced by cyanobacteria (Winston et al., 2014). These relationships are concerning considering that, with changes in our climate, heat waves and droughts may be increasing in the U.S. (Petersen et al., 2014). The complex phytoplankton community in Beaver Lake is likely influenced by both nutrient inputs as well as the hydro-climate, but further research is needed to fully characterize what taxa are present and how the community reacts to changes in the reservoir.

Conclusions

Beaver Lake is at risk of violating its chl-a standard (Scott and Haggard, 2015). Increases in TN as well as chl-a at the Lowell site are of concern for water management and sources of nutrients need to be understood. Trend analysis of the tributaries from 2009 to 2015 revealed that SRP has decreased in War Eagle Creek while SRP and TP increased at Richland Creek. Monotonic trends imply increases in TN at Richland Creek and White River; however, NO_3 and TN concentrations throughout the UWRB reflect an increase from 2009 to late 2012/2013, and either declining or maintaining through the later period. Large variability in the hydro-climate is likely a major driver in N concentrations as well as algal productivity, where hot and dry periods may increase baseflow N concentrations and elevate algal growth. The assumption of stationarity (i.e., stochastic systems exhibits statistical properties that do not change with time) is considered to be no longer valid (Milly et al., 2008), thus, real persistent trends in monitoring data are difficult to separate from large-scale variability (Cohn and Lins, 2005). In order to better characterize how the UWRB is changing with time, long-term (i.e., on the order of decades) high-quality data (i.e., sufficient coverage across the seasonal, hydrological, and temporal records) are needed (Burt et al., 2014). Short-term records can characterize some important events (e.g., certain hot biogeochemical processes); long-term records can inform strategic planning across the watershed.

Table 1. Trend magnitudes, in percent change in concentration per year (% change y^{-1}), for nitrate-nitrogen (NO_3), total nitrogen (TN), soluble reactive phosphorus (SRP), and total phosphorus (TP) at the three inflows into Beaver Lake from 2009 to 2015, where trends were estimated using just baseflow data, just storm data, and all data combined; only significant trends (via overall F-test; $p<0.10$) are reported.

		Trend in Constituent Concentration (% change y^{-1})			
Site	**Flow Regime**	**NO_3**	**TN**	**SRP**	**TP**
Richland Creek	Baseflow	11.3	8.3	-	9.1
	Storm	-	5.2	-	-
	All	8.4	7.3	5.2	7.4
War Eagle Creek	Baseflow	-	-	-	-
	Storm	-6.2	-	-11.7	-
	All	-	-	-3.7	-
White River	Baseflow	-	6.3	-	-
	Storm	-	-	-	-
	All	-	3.8	-	-

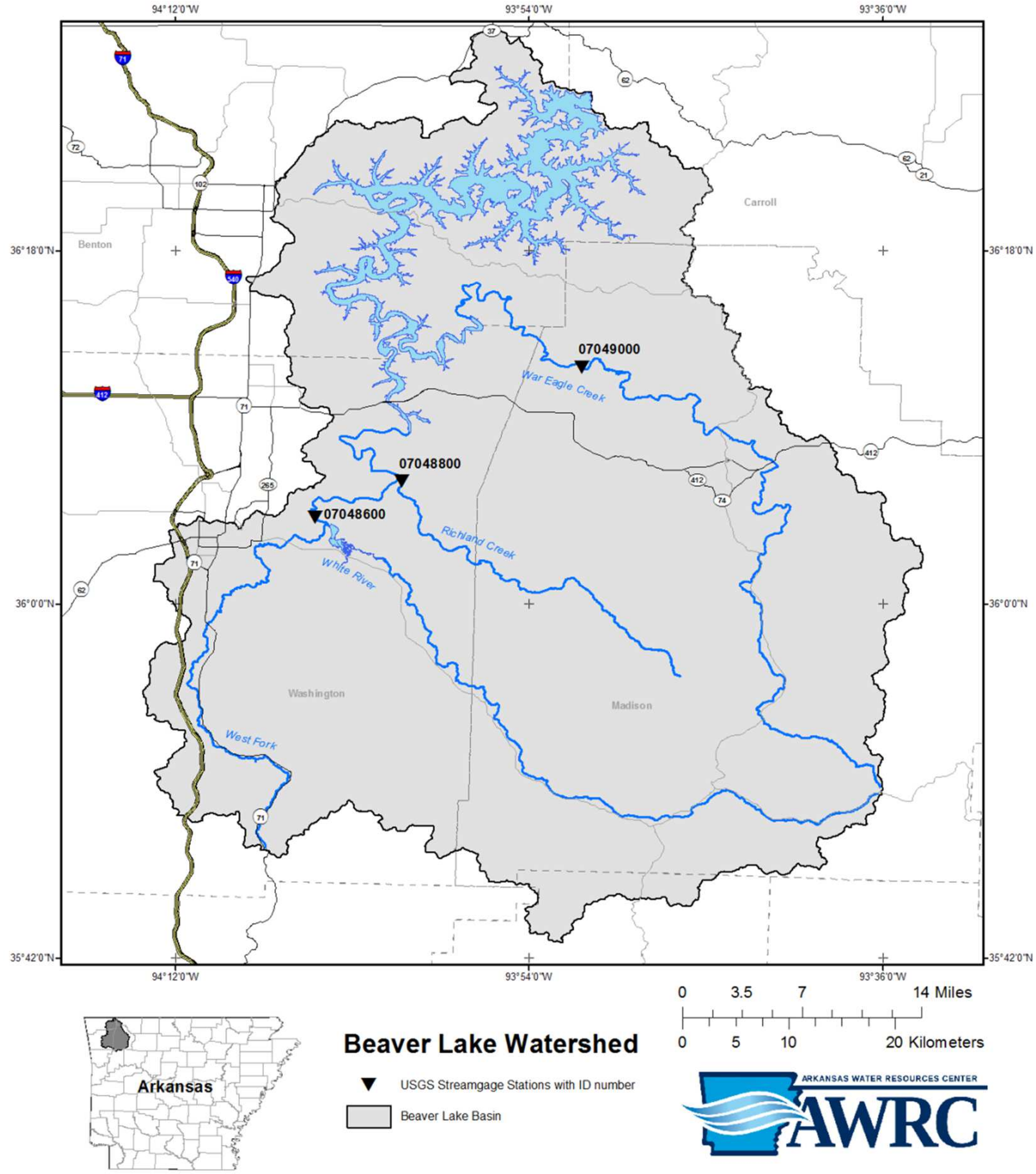

Figure 1. Map of the Upper White River Basin within the Beaver Lake Watershed with sampling sites at White River (USGS 07048600), Richland Creek (07048800), and War Eagle Creek (07049000).

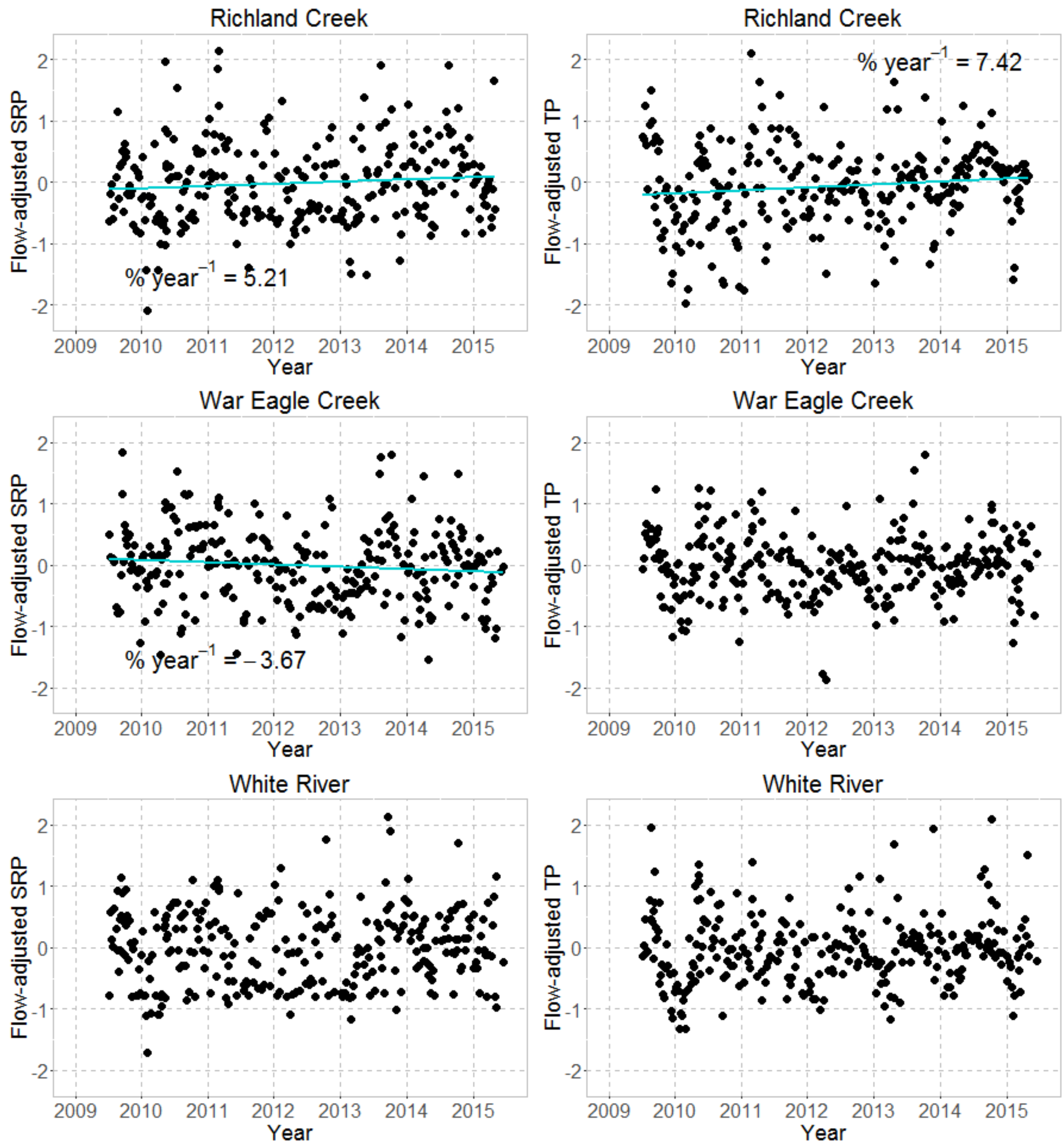

Figure 2. Trend analyses of soluble reactive phosphorus (SRP) and total phosphorus (TP) under all flow conditions at Richland Creek, War Eagle Creek, and White River from 2009 through 2015; significant monotonic trends (F-test, $p<0.10$) are shown with a green line with magnitude given in percent change in concentration per year.

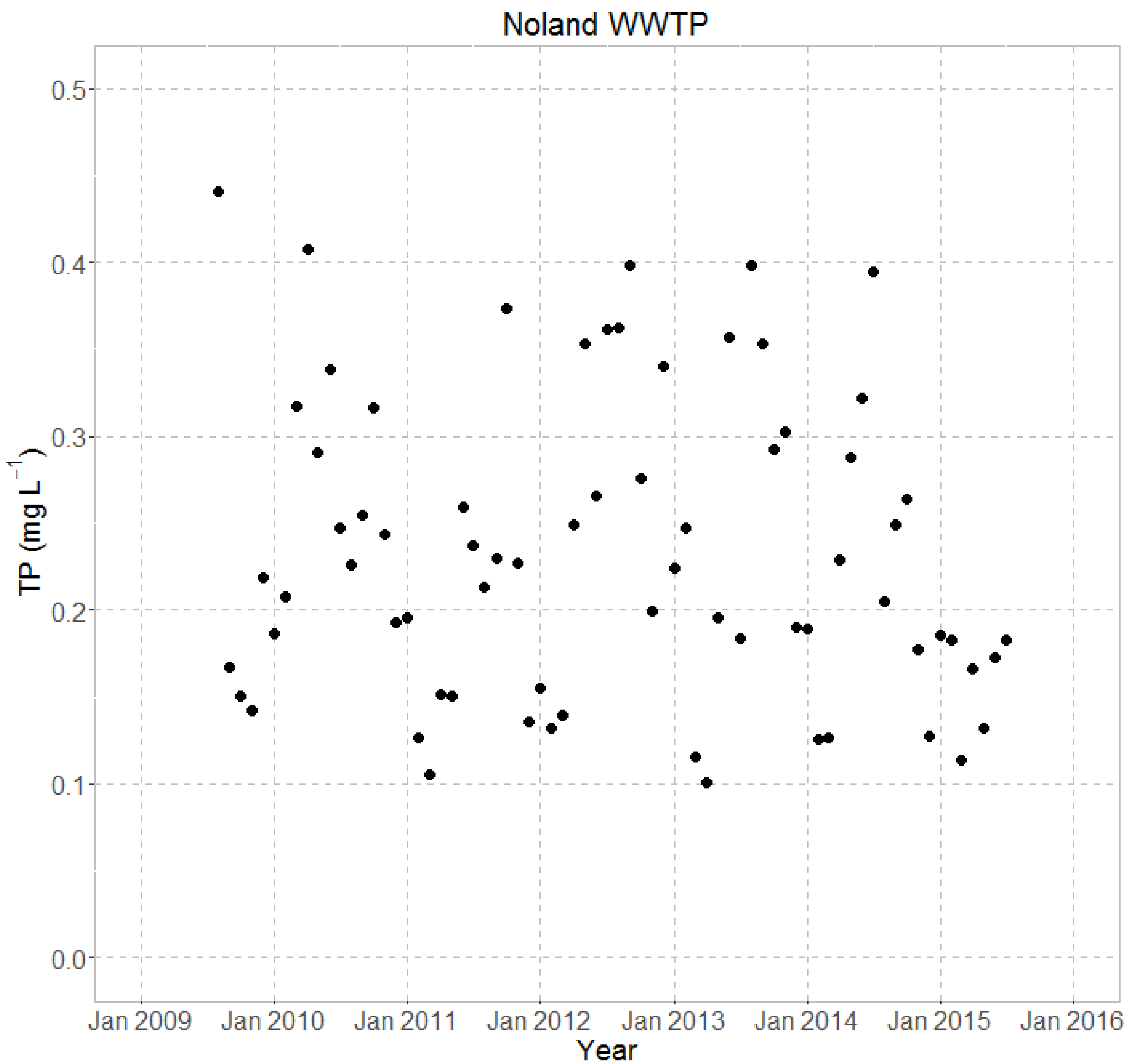

Figure 3. Monthly average effluent total phosphorus (TP) concentrations at the Paul Noland wastewater treatment plant (WWTP) in Fayetteville, Arkansas; this facility discharges into the White River downstream from the site in our study.

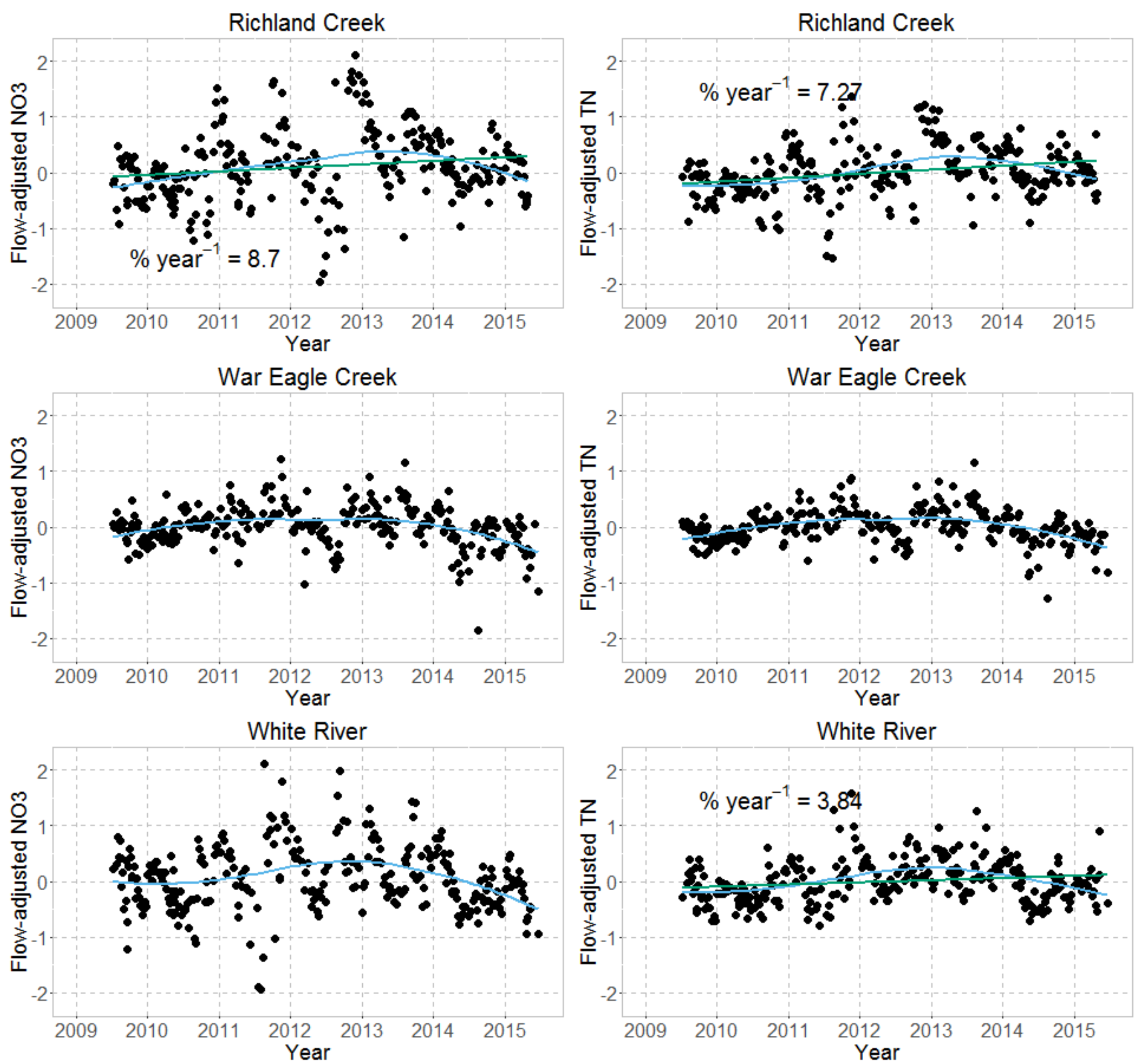

Figure 4. Trend analyses of nitrate-nitrogen (NO_3) and total nitrogen (TN) under all flow conditions at Richland Creek, War Eagle Creek, and White River from 2009 through 2015; significant monotonic trends (F-test, $p<0.10$) are shown with a green line with magnitude given in percent change in concentration per year, and a smoother (in blue) is fit to the data to visualize subtle changes over time.

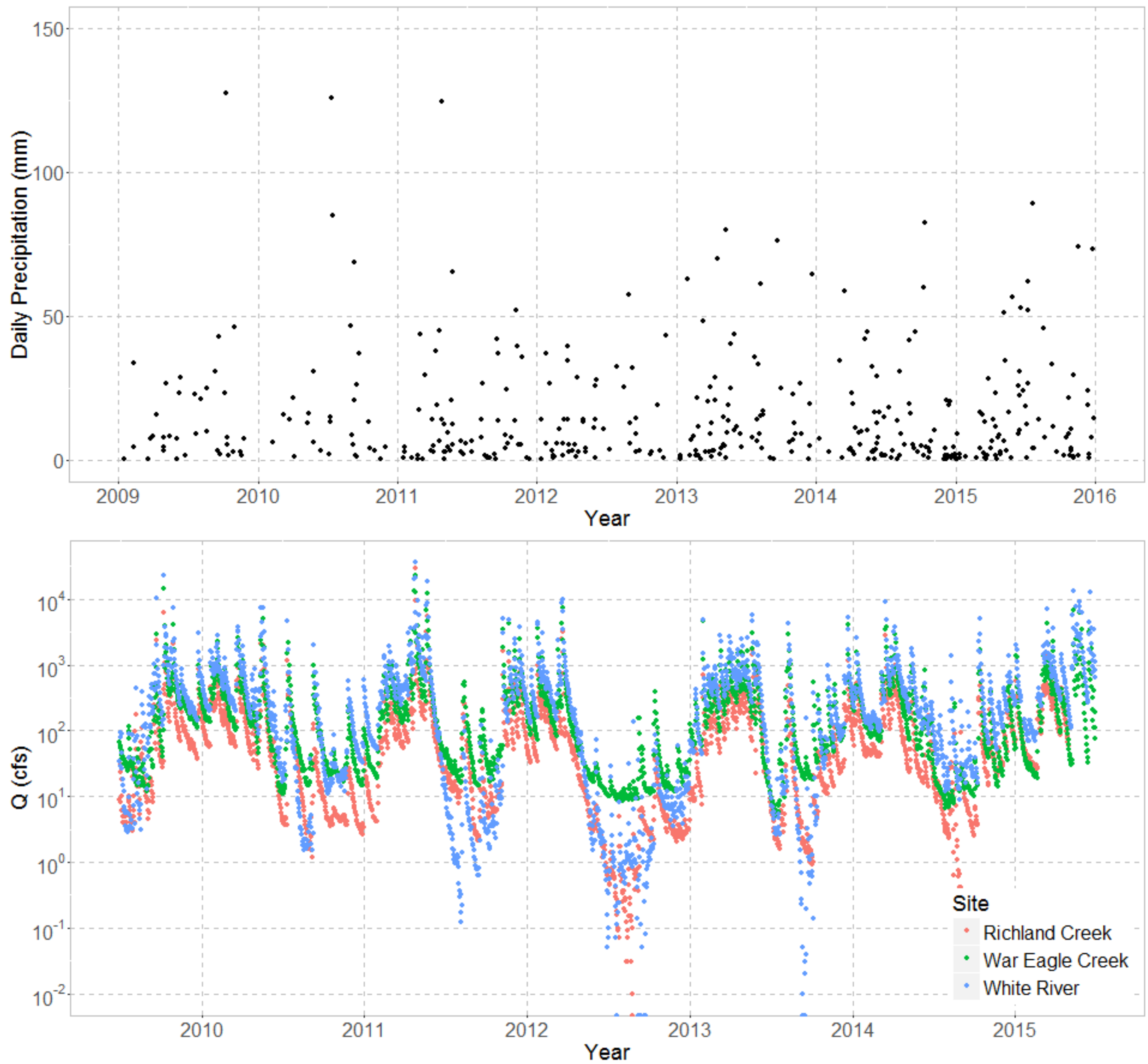

Figure 5. (Top) Daily precipitation (mm) measured in Fayetteville, Arkansas from 2009 through 2015 (NADP, 2016), and (Bottom) daily mean stream flow (Q; cfs) measured at Richland Creek, War Eagle Creek, and White River from July 2009 through June 2015.

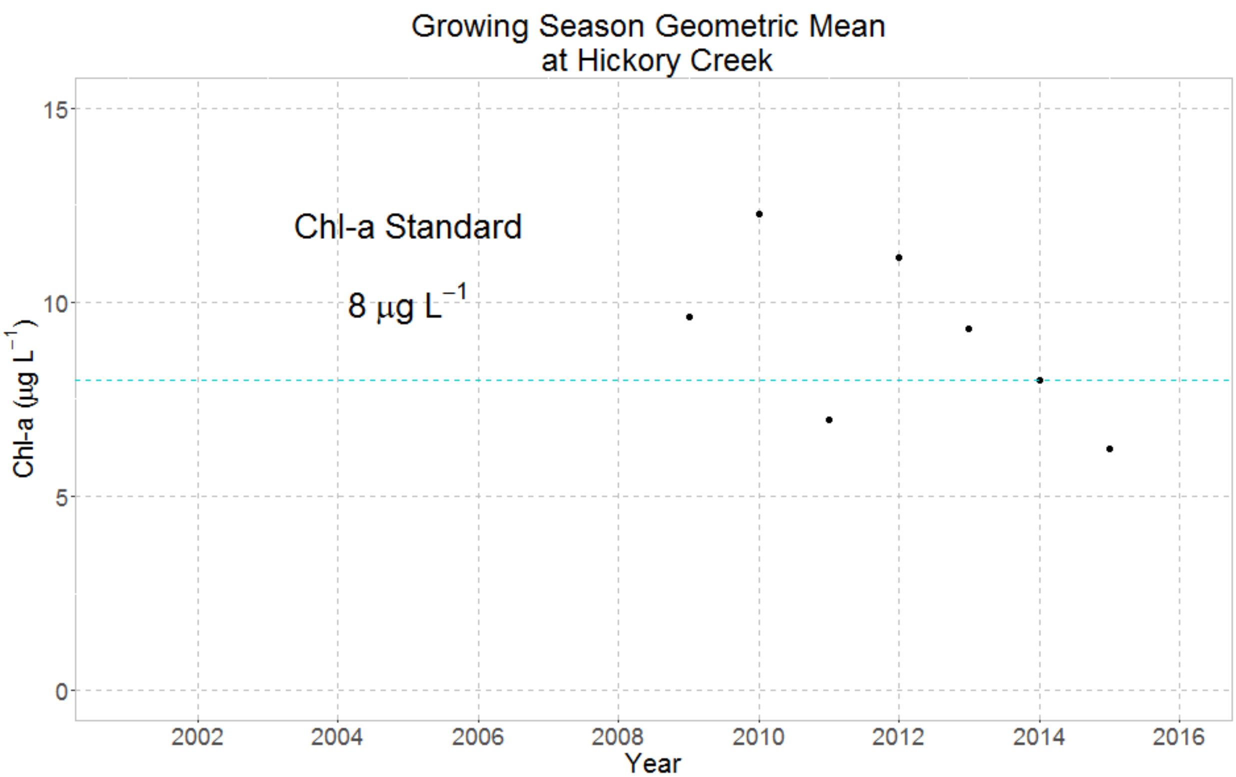

Figure 6. Growing season (May through October) geometric mean chlorophyll-a in Beaver Lake at Hickory Creek, and the water quality standard for Beaver Lake is shown (dashed line).

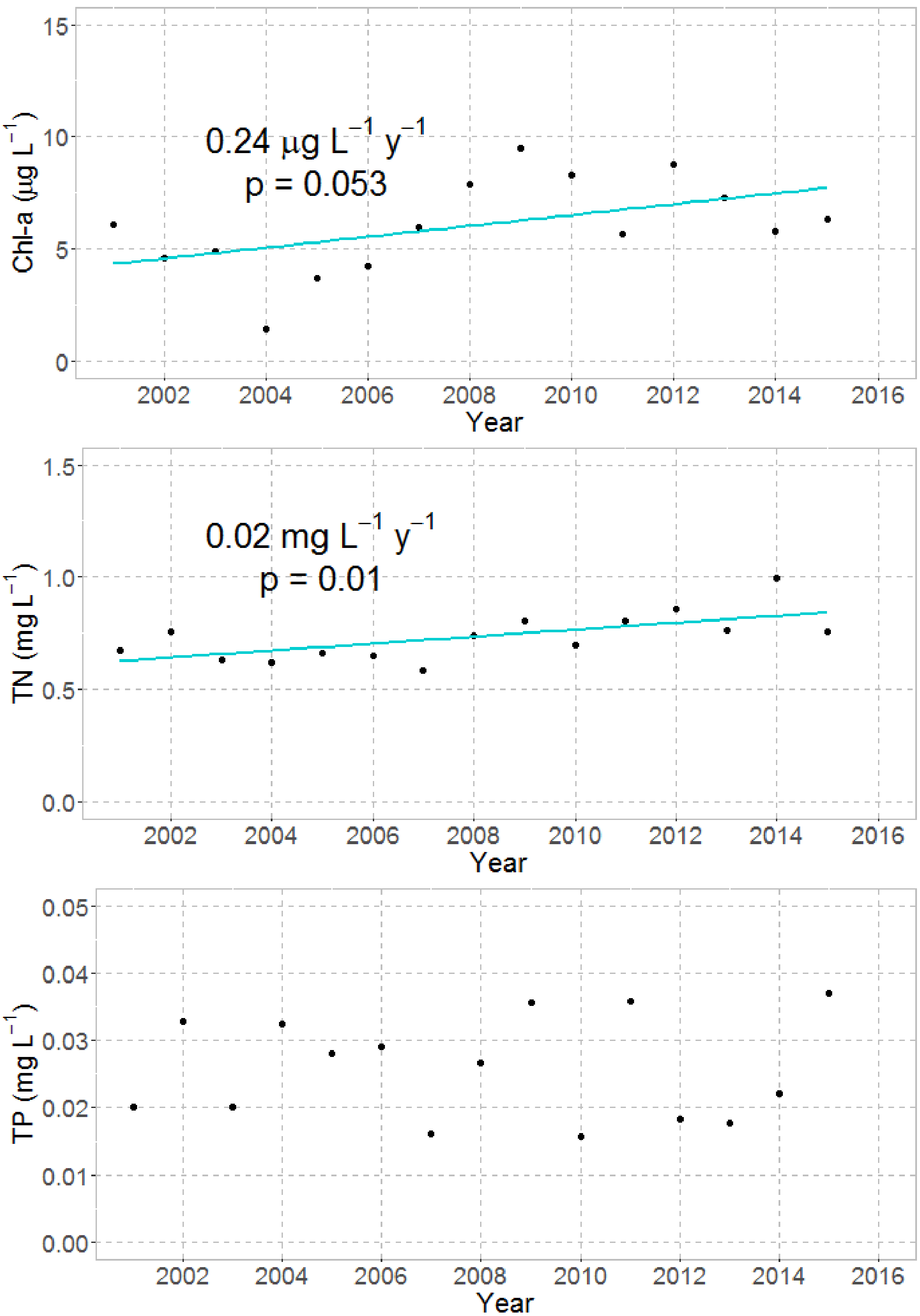

Figure 7. Growing season (May through October) geometric means for chlorophyll-a (chl-a) and annual average total nitrogen (TN) and total phosphorus (TP) concentrations in Beaver Lake at Lowell, Arkansas; linear regression trend lines are given (blue line, $p<0.10$).

Works Cited

Arkansas Department of Environmental Quality (ADEQ). 2016. Assessment Methodology – For the preparation of The 2016 Integrated Water Quality Monitoring and Assessment Report. Accessed at https://www.adeq.state.ar.us/water/planning/integrated/303d/pdfs/2016/assessment-methodology.pdf

Arkansas Pollution Control and Ecology Commission (APCEC). 2015. Regulation No. 2: Regulation establishing water quality standards for surface waters of the state of Arkansas. Accessed at https://www.adeq.state.ar.us/regs/files/reg02_final_151114.pdf

AWRC. 2016. AWRC Water Quality Laboratory. Accessed at http://arkansas-water-center.uark.edu/waterqualitylab.php on April 21, 2016.

Bailey, B.W., B.E. Haggard, and L.B. Massey. 2012. Water quality trends across select 319 monitoring sites in Northwest Arkansas. Arkansas Water Resources Center, Fayetteville, Arkansas, MSC Publication 365: 50 pp.

Beaver Water District. 2014. Fact Sheet: Stage 2 Disinfectants and Disinfection Byproduct Rule. Accessed at http://www.bwdh2o.org/wp-content/uploads/2014/05/DBPs-Fact-Sheet-UPDATE-2014.pdf

Bowes, M.J., H.P. Jarvie, S.J. Halliday, R.A. Skeffington, A.J. Wade, M. Loewenthal, E. Gozzard, J.R. Newman, and E.J. Palmer-Felgate. 2015. Characterising phosphorus and nitrate inputs to a rural river using high-frequency concentration-flow relationships. Science of the Total Environment 511: 608-620.

Boyer, D.G. and G.C. Pasquarell. 1995. Nitrate concentrations in karst springs in an extensively grazed area. American Water Resources Association 31(4): 729–736.

Brown, T.C. and P. Froemke. 2012. Nationwide assessment of nonpoint source threats to water quality. BioScience, 62: 136-146.

Burt, T.P., N.J.K. Howden, and F. Worrall. 2014. On the importance of very long-term water quality records. Wiley Interdisciplinary Reviews: Water, 1(1): 41-48.

CAST. 2006. Arkansas Watershed Information System: a module of the Arkansas Automated Reporting and Mapping System. Accessed at: http://watersheds.cast.uark.edu/index.php

Chapman, T.G. 1991. Comment on “Evaluation of Automated Techniques for Base Flow and Recession Analyses” by R.J. Nathan and T.A. McMahon. Water Resources Research 27(7): 1783-1784.

Cohn, T.A. and H.F. Lins. 2005. Nature’s style: naturally trendy. Geophysical Research Letters 32(23).

Cooke, G.D., E.B. Welch, and J.R. Jones. 2011. Eutrophication of Tenkiller Reservoir, Oklahoma, from nonpoint agricultural runoff. Lake and Reservoir Management 27: 256-270.

DeLaune, P.B., P.A. Moore Jr., D.K. Carman, A.N. Sharpley, B.E. Haggard, and T.C. Daniel. 2004. Development of a phosphorus index for pastures fertilized with poultry litter - factors affecting phosphorus runoff. Journal of Environmental Quality 33:2183-2191.

Dodds, W.K., W.W. Bouska, J.L. Eitzmann, T.J. Pilger, K.L. Pitts, A.J. Riley, J.T. Schloesser, and D.J. Thornbrugh. 2009. Eutrophication of U.S. Freshwaters: Analysis of Potential Economic Damages. Environmental Science and Technology 43(1): 12-19.

Eckhardt, K. 2005. How to construct recursive digital filters for baseflow separation. Hydrological Processes 19: 507-515.

Eckhardt, K. 2008. A comparison of baseflow indices, which were calculated with seven different baseflow separation methods. Journal of Hydrology 352: 168-173.

Finlay, J.C., G.E. Small, and R.W. Sterner. 2013. Human influences on nitrogen removal in lakes. Science 342: 247-250.

Giovannetti, J., L.B. Massey, B.E. Haggard, and R.A. Morgan. 2013. Land use effects on stream nutrients at Beaver Lake Watershed. Journal of American Water Works Association 105(1): E1-E10.

Green, W.R., and B.E. Haggard. 2001. Phosphorus and nitrogen concentrations and loads at Illinois River South of Siloam Springs, Arkansas, 1997-1999. U.S. Geological Survey. Water-Resources Investigations Report 01-4217.

Haggard, B.E., P.A. Moore Jr., I. Chaubey, and E.H. Stanley. 2003. Nitrogen and phosphorus concentrations and export from an Ozark Plateau catchment in the United States. Biosystems Engineering 86(1): 75-85.

Helsel, D.R., and R.M. Hirsch. 2002. Statistical methods in water resources. In *Hydrologic Analysis and Interpretation,* (pp. 1-510). United States Geological Survey.

Hirsch, R.M., R.B. Alexander, and R.A. Smith. 1991. Selection of methods for the detection and estimation of trends in water quality. Water Resources Research 27(5): 803-813.

Hufhines, K.R. Brye, B.E. Haggard, and R. Morgan. 2011. Net nutrient uptake in the White River, Northwest Arkansas, downstream of a municipal wastewater treatment plant. Journal of Environmental Protection 2: 255-270.

Jarvie, H.P., A.N. Sharpley, J.T. Scott, B.E. Haggard, M.J. Bowes, and L.B. Massey. 2012. Within-river phosphorus retention: accounting for a missing piece in the watershed phosphorus puzzle. Environmental Science and Technology 46: 13284-13292.

Jarvie, H.P., A.N. Sharpley, V. Brahana, T. Simmons, A. Price, C. Neal, A.J. Lawlor, D. Sleep, S. Thacker, and B.E. Haggard. 2014. Phosphorus retention and remobilization along hydrological pathways in karst terrain. Environmental Science and Technology 48(9): 4860-4868.

McDowell, R.W., B.J.F. Biggs, A.N Sharpley, and L. Nguyen. 2004. Connecting phosphorus loss from agricultural landscapes to surface water quality. Chemistry and Ecology 20(1): 1 -40.

Migliaccio, K.W., B.E. Haggard, I. Chaubey, and M.D. Matlock. 2007. Linking watershed subbasin characteristics to water quality parameters in War Eagle Creek watershed. Transactions of the American Society of Agricultural and Biological Engineers 50(6): 2007-2016.

Mosley, L.M. 2015. Drought impacts on the water quality of freshwater systems; review and integration. Earth-Science Reviews 140: 203-214.

Nathan, R.J. and T.A. McMahon. 1990. Evaluation of automated techniques for base flow and recession analyses. Water Resources Research 26(7): 1465-1473.

National Atmospheric Deposition Program (NADP). 2016. National Trends Network Site AR27. Accessed at http://nadp.sws.uiuc.edu/data/sites/sitedetails.aspx?id=AR27&net=NTN

National Oceanic and Atmospheric Administration. 2016. National Climatic Data Center: Climate Indices. Accessed at http://www7.ncdc.noaa.gov/CDO/CDODivisionalSelect.jsp#

Nürnberg, G.K. 1996. Trophic state of clear and colored, soft- and hardwater lakes with special consideration of nutrients, anoxia, phytoplankton and fish. Lake and Reservoir Management 12(4): 432-447.

Paerl, H.W. and J. Huisman. Blooms like it hot. Science 320: 57-58.

Scott, J.T., J.K. Stanley, R.D. Doyle, M.G. Forbes, and B.W. Brooks. 2009. River-reservoir transition zones are nitrogen fixation hot spots regardless of ecosystem trophic state. Hydrobiologia 625(1): 61-68.

Scott, J.T. and B.E. Haggard. 2015. Implementing effects-based water quality criteria for eutrophication in Beaver Lake, Arkansas: Linking standard development and assessment methodology. Journal of Environmental Quality 44(5): 1503-1512.

Scott, J.T. and M.J. McCarthy. Nitrogen fixation may not balance the nitrogen pool in lakes over timescales relevant to eutrophication management. Limnology and Oceanography 55(3): 1265-1270.

Sen, S., B.E. Haggard, I. Chaubey, K.R. Brye, T.A. Costello, and M.D. Matlock. 2007. Sediment phosphorus release at Beaver Reservoir, Northwest Arkansas, USA, 2002-2003: A preliminary investigation. Water, Air, and Soil Pollution 179: 67-77.

Simpson, Z.P. and B.E. Haggard. 2016A. Are phosphorus concentrations still declining in the Illinois River Watershed, Northwest Arkansas? In preparation.

Simpson, Z.P. and B.E. Haggard. 2016B. An optimized procedure for flow-adjustment of constituent concentrations for trend analysis. In preparation.

Stamm, C., H.P. Jarvie, and J.T. Scott. What's more important for managing phosphorus: loads, concentrations or both? Environmental Science and Technology 48: 23-24.

Sterner, R. 2008. Review Paper: on the phosphorus limitation paradigm for lakes. International Review of Hydrobiology 93(4-5): 433-445.

Walker Jr., W.W. 1983. Significance of eutrophication in water supply reservoirs. Journal of American Water Works Association 1983: 38-42.

Winston, B., S. Hausmann, J.T. Scott, and R. Morgan. 2014. The influence of rainfall on taste and odor production in a south-central USA reservoir. Freshwater Science 33(3): 755-764.

Conclusion

Trend analysis is a valuable tool for water resources management, but should be wary of how a constituent covaries with streamflow. The first chapter of this thesis developed an improvement to a flow-adjustment procedure by optimizing the smoothing parameter in LOESS. It is recommended to use 10 x 10 cross-validation to determine an optimal fit to the data. While there was no significant difference in trend interpretations for the datasets used in this study, the proposed method may have benefits with other datasets containing more/less data, longer period of record, or differently-behaving constituents.

The case study of phosphorus (P) in the Illinois River Watershed (IRW) in the second chapter of this thesis found significant decreases in the dissolved form of P at three sites located downstream from major wastewater treatment plants (WWTPs). However, total P (TP) concentrations have not continued to decrease at the Illinois River. Legacy sources of P in the IRW may be buffering the positive effects of watershed management, prolonging the period until desired changes take place. Future studies could examine the mechanisms behind legacy P sources; meanwhile, it is important for watershed managers to remain determined.

The last chapter of this thesis, the case study of nutrient trends in the inflows to Beaver Lake, illustrated the importance of long-term monitoring. Particularly, nitrogen (N) concentrations showed a similar pattern for the study period across each sampling site, where flow-adjusted concentrations increased from 2009 to 2013 then declined or maintained for the remainder of the study period. This correlated with a major drought that occurred in 2012, providing a disturbance in the hydro-climate across this watershed. Elevated chlorophyll-a concentrations were found at both sites in Beaver Lake, and the site near a drinking water intake exhibited increases in total N and chlorophyll-a from 2001 to 2015. It will be important for future studies to use long-term data of the inflows to understand processes driving nutrient loadings and determine whether persistent trends exist in the data.