Illinois River Watershed Total Phosphorus Criterion Revision

Oklahoma Water Resources Board
3800 North Classen Blvd
Oklahoma City, OK 73118

Staff Report
December 1, 2020

Contents

List of Figures

List of Tables

Introduction

The waters within the Illinois River watershed are among Oklahoma's most beautiful and popular waters. In fact, the three primary rivers within the watershed are all Oklahoma scenic rivers and are recognized to have great aesthetic, ecological, and recreational value. The protection of these waters by means of water quality standards is of the upmost importance to the state of Oklahoma. Water quality standards (WQS) define the goals for a waterbody and work to safeguard human health and aquatic life by establishing provisions to limit pollution to the state's lakes, rivers, and wetlands. Water quality standards are comprised of three components 1) a waterbody's beneficial uses, 2) water quality criteria, and 3) the antidegradation policy. Beneficial uses establish the water quality goals for the waterbody, criteria define the minimum water quality condition necessary to achieve those goals, and the antidegradation policy specifies the framework to be used in making decisions regarding any intentional lowering of water quality. The antidegradation policy ensures that good water quality is conserved where possible and lowered only when necessary, that stakeholders affected by the lowering are included in the process, and that beneficial uses are maintained and protected.

The Oklahoma Water Resources Board (OWRB) is the state agency responsible for promulgating water quality standards to ensure water quality protection across the state (82 O.S. §1085.30). Oklahoma has long recognized the importance of maintaining and protecting the state's waters through adoption of water quality standards. The water quality standards are set forth in Oklahoma Administrative Code Title 785, Chapter 45 and the Implementation of Oklahoma's Water Quality Standards are set forth in Title 785, Chapter 46. Consistent with both Oklahoma and federal regulations (785:45-1-1, 40 CFR § 131.11) all water quality criteria must be established to ensure protection of beneficial uses. This document presents the technical background information used by the OWRB in developing changes to the total phosphorus criterion that protects the Aesthetic beneficial uses of waters in the Illinois River Watershed. The total phosphorus criterion applicable to the Mountain Fork River, Lee Creek and Little Creek was not subject to any revision and is not addressed in this document.

Environmental Setting and Background

The Illinois River watershed (HUC 11110103) is located in northeastern Oklahoma and northwestern Arkansas and spans the political boundary between the two states. The watershed area is about 1,654 square miles (Figure 1). The mainstem of the Illinois River originates in the

Boston Mountains in Washington County, Arkansas. The river flows north for approximately 36 miles and turns westward at the confluence with Osage Creek; from here it flows west into Oklahoma. Flint Creek is a major tributary to the Illinois River. Flint Creek drains 127 square miles in the northwest portion of the watershed and has its confluence with the Illinois River just south of the Oklahoma state highway 59. Below this confluence the Illinois River flows southwest past the city of Tahlequah to Tenkiller Ferry Reservoir (Lake Tenkiller). Barren Fork Creek is another major tributary that has a confluence with the Illinois River just before it enters Lake Tenkiller. Barren Fork Creek drains 346 square miles in the central area of the watershed. Below the Lake Tenkiller Dam the Illinois River flows 9.5 miles to its confluence with the Arkansas River.

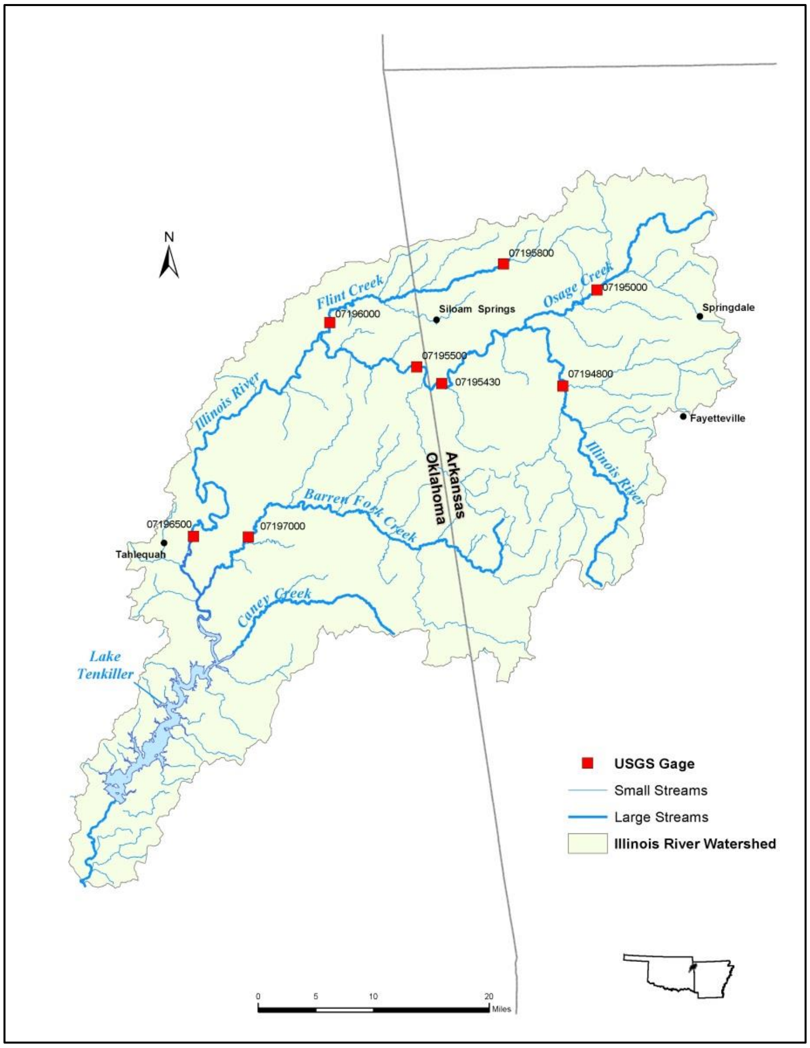


Figure 1 Illinois River Watershed map with USGS stream gages.

There are a number of United States Geological Service (USGS) gages in the Illinois River watershed; those listed in Table 1 were used for analysis throughout this report (Lewis et al., 2009 and Granato et al., 2017).

Table 1 USGS Stream gages relied upon throughout this report.

Location Name	Gage Station Number	Flow Record Used in this Analysis		Tributary Area (mi^2)	Average Daily Streamflow (cfs)	Daily Flow Years
Osage Creek near Elm Springs, AR	07195000	Jan. 2008	Dec.2018	130	123	34
Illinois River at Savoy, AR	07194800	Jan. 2008	Dec.2018	167	145	12
Illinois River at South of Siloam Springs, AR	07195430	Jan. 2008	Dec.2018	575	573	9
Illinois River near Watts, OK	07195500	Jan. 2008	Dec.2018	630	621	52
Flint Creek near Kansas, OK	07196000	Jan. 2008	Dec.2018	116	116	50
Illinois River near Tahlequah, OK	07196500	Jan. 2008	Dec.2018	950	929	72
Barren Fork at Eldon, OK	07197000	Jan. 2008	Dec.2018	312	325	59

Based on the 2016 National Land Cover Database, forty-one percent of the land use in the Illinois River watershed is classified as deciduous forest and almost forty percent is hay and pasture (Table 2, Figure 2). These two land uses alone dominate the watershed landscape. Developed areas (open space, low, medium, and high intensity) only account for about ten percent of the watershed. The remaining ten percent of the watershed is mostly open water and grass/shrub/forest areas.

Table 2 NLDC Land Use for 2016

Land Use Class	Area (square miles)	Percentage (%)
Deciduous Forest	678.3	41.0
Hay/Pasture	655.7	39.7
Developed, Open Space	91.4	5.5
Mixed Forest	66.7	4.0
Developed, Low Intensity	46.0	2.8
Open Water	26.9	1.6
Developed, Medium Intensity	24.4	1.5
Shrub/Scrub	17.7	1.1
Herbaceous	16.7	1.0
Evergreen Forest	9.9	0.6
Developed, High Intensity	9.4	0.6
Woody Wetlands	6.5	0.4
Cultivated Crops	1.9	0.1
Barren Land	1.7	0.1
Emergent Herbaceous Wetlands	0.5	0.03
Total of classes	**1653.8**	**100.0**

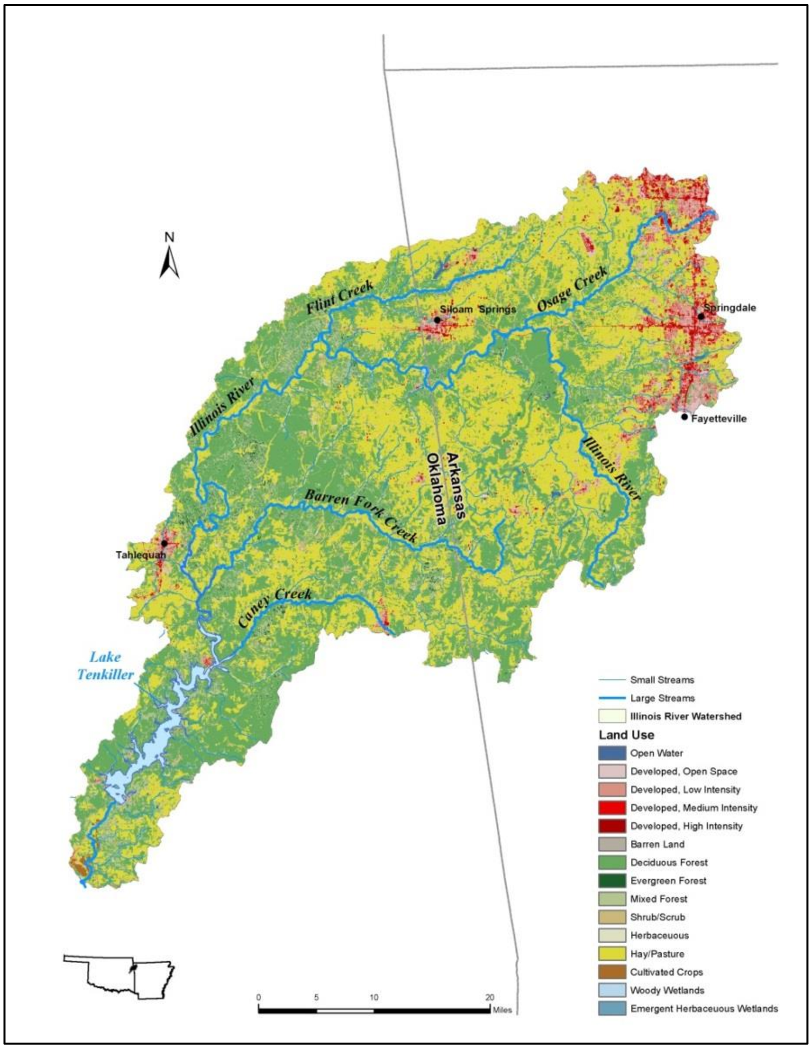


Figure 2 NLDC Land Use for 2016

The Illinois River watershed and Lake Tenkiller are among Oklahoma’s most beautiful and popular waters. The Oklahoma WQS designated beneficial uses for these waters are public water supply, aquatic life, aesthetics, body contact recreation, and agriculture. There are millions of visitors to these waterbodies annually who enjoy activities like swimming, boating, fishing, and scenic vistas. These visitors contribute substantially to the local economic activity; for example, in 2019 visitor spending within 30 miles of Lake Tenkiller was just under $70 million dollars (USACE 2019). In addition to monetary measures of economic activity, the experiences people enjoy on the Illinois River are equally valuable. The Oklahoma legislature recognized the intrinsic value of the Illinois River, Flint Creek, and Barren Fork Creek with the 1970 Scenic Rivers Act (82 O.S. 1451-1471). This act states:

> *“The Oklahoma Legislature finds that some of the free-flowing streams and rivers of Oklahoma possess such unique natural scenic beauty, water conservation, fish, wildlife and outdoor recreational values of present and future benefit to the people of the state that it is the policy of the Legislature to preserve these areas for the benefit of the people of Oklahoma. For this purpose there are hereby designated certain "scenic river areas" to be preserved as a part of Oklahoma's diminishing resource of free-flowing rivers and streams.”*

Grounded in this statue and Oklahoma’s declaration that scenic rivers shall be preserved for the benefit of the people, rules were adopted to protect water quality. In particular, the WQS Antidegradation Policy, in place since 1973, classifies the scenic rivers as Outstanding Resources Waters (ORW) and affords to them additional protections (see Antidegradation Policy section).

In 2001 it was recognized that Oklahoma’s scenic rivers were being negatively impacted by phosphorus pollution and in response the OWRB adopted a total phosphorus criterion to protect the aesthetic beneficial use of these waters. The adopted criterion is presented below (785:45-5-19(c)(2)).

> *“The thirty (30) day geometric mean total phosphorus concentration in waters designated Scenic River…shall not exceed 0.037 mg/L.”*

The scientific basis for this criterion relied up the reference approach outlined in the EPA *Nutrient Criteria Technical Guidance Manual for Rivers and Streams* (EPA, 2000) where an upper percentile of the total phosphorus concentration distribution was used to set the criterion magnitude. Distributions of flow-weighted total phosphorus concentrations from 85 relatively unimpacted streams across the United States were presented in the paper *Nutrient Concentrations and Yield in Undeveloped Stream Basins of the United States* (Clark et. al., 2000); the 75th percentile phosphorus concentration of 0.037 mg/L from this analysis was employed as the criterion magnitude. This phosphorus value was reviewed for relevance in Oklahoma waterbodies and found to be similar to TP concentrations observed in the Mountain Fork River (also a scenic river) of 0.028 mg/L and the comparatively less disturbed Barren Fork River of 0.045 mg/L (OWRB, 2012).

As a consequence of the criterion adoption by OWRB, the states of Oklahoma and Arkansas entered into a *Statement of Joint Principles and Actions* under which the states agreed to coordinated, but independent actions to reduce phosphorus loading in the Illinois River watershed. A review of actions taken and progress made on phosphorus reduction under this agreement is beyond the scope of this document. More information regarding this agreement may be obtained from partner agencies such as Oklahoma Conservation Commission, Oklahoma Department of Environmental Quality, Arkansas Division of Environmental Quality (ADEQ), and Arkansas Department of Agriculture, Natural Resources Division. Under the *Statement of Joint Principles and Actions,* Oklahoma agreed to a reevaluation of the adopted criterion in 2012 based on the best scientific information available and with participation from Arkansas.

In 2011-2012 OWRB staff conducted a criterion reevaluation project in collaboration with Oklahoma environmental agencies, the Cherokee Nation, ADEQ, Arkansas Natural Resources Commission (now Natural Resources Division), and EPA Region 6. Representatives from these organizations were convened as a Technical Advisory Group (TAG) for the purpose of reviewing the best scientific information available and to provide a recommendation to OWRB staff if actions to revise the total phosphorus criterion were warranted (OWRB, 2012). A literature review was conducted under a TAG reviewed and EPA approved secondary Quality Assurance Project Plan (QAPP) (OWRB, 2011). The QAPP stipulated key literature review components such as subject matter, acquisition methods, rankings for information quality and geographic relevance, decision rules, and potential decision recommendations. In total, 136 technical studies were reviewed under the reevaluation and summaries of the 10 most relevant studies were presented in the *Oklahoma Scenic Rivers Phosphorus Criteria Review, Final Report.* The median total phosphorus concentration recommended from these studies was 0.036 mg/L and the mode was 0.020 mg/L. In accordance with decision rules set in the QAPP, Oklahoma's total phosphorus criterion magnitude was found to be aligned with the best scientific information available and the TAG did not recommend a criterion revision.

Arkansas members of the TAG however, did not concur with the overall TAG recommendation and submitted a minority report to the OWRB (Arkansas TAG, 2012). The minority report stated that the literature review did not substantiate the total phosphorus criterion as the concentration necessary or appropriate to protect Oklahoma's scenic rivers and believed that there was a significant lack of measured data, particularly on the Illinois River mainstem within Oklahoma state boundaries. This report included four recommendations to the OWRB, one of which was to conduct a stressor response study on the Illinois River. In 2013 the states of Oklahoma and Arkansas entered into the *Second Statement of Joint Principles and Actions*.

The *Second Statement of Joint Principles and Actions* between Arkansas and Oklahoma environmental agencies was signed in February 2013. Under this agreement the states completed the Joint Phosphorus Criteria Study (Joint Study) managed by the Joint Study Committee. The committee was composed of six members, three from each state. The Joint Study was conducted from 2014 through 2016 and culminated in a Final Report submitted to both state governors on December 19, 2016. The key task of the Joint Study Committee was to make a recommendation regarding *"… what phosphorus levels, and what frequency and duration components of measure, are necessary to protect the aesthetics beneficial use and scenic river (Outstanding Resource Water) designations…"* To that end, the Joint Study Committee made the following recommendation in the Final Report (Joint Study, 2016).

> *"A six-month average total phosphorus level of not to exceed 0.035 mg/L based on water samples taken during the critical condition…"*

OWRB staff has conducted this criterion revision as an outgrowth of recommendations from the Arkansas Oklahoma Joint Study Committee.

Problem Identification

Nutrients, including nitrogen (N) and phosphorus (P), are essential for plant growth and are often important limiting nutrients in aquatic environments. However, in situations of nutrient enrichment, the nutrients N and P are no longer limiting; in fact, they are readily available in the waterbody, which causes an increase in primary production and eutrophication. Eutrophication is defined by increased nutrient loading to a waterbody and the subsequent ecological response. Abundant input of nutrients into rivers leads to degraded waterbody conditions. Symptoms of eutrophication in rivers are listed below.

- Increased algal biomass (macroalgae and phytoplankton)
- Reductions in dissolved oxygen (hypoxia)
- Alterations in algal species composition
- Alterations in food resources and habitat structure
- Harmful algal blooms

The relationship between nuisance algae growth and nutrient enrichment in stream systems has been well-documented in the literature (Dodds and Welch, 2000; Biggs, 2000; Busse et al., 2006). Eutrophication and nutrient enrichment problems rank as one of the top causes of water quality impairment; phosphorus and nitrogen are the most widespread chemical stressor to the nation's waters (EPA, 2017). The problems associated with these impacts can range from a recreational nuisance to serious aquatic life and public health concerns. For example, high amounts of algal biomass and other aquatic plants interfere with swimming or wading, angling, and/or aesthetic enjoyment of the waterbody and impair the recreational beneficial uses. The aquatic life impacts of eutrophication can include fish kills, lowered fishery production, loss or degradation of important habitats (e.g. cobble/gravel niche space), and smothering of benthic organisms (EPA CADDIS).

There are many complex ways in which excess nutrient loads can impact beneficial uses. The conceptual model in Figure 3 outlines the interactions between nutrients and biological responses in streams. There are numerous overlapping physical, chemical, and biological factors that affect how a waterbody responds to increased nutrient loading. For example, nutrients, temperature, and light often interact together and influence processes within the aquatic ecosystem. The model below demonstrates the interaction and influence of various factors and to assess pathways that are contributing to the impairment of beneficial uses.

Increased nutrient loading into the stream can result in increased algal growth (Figure 3). The high levels of algal biomass through respiration (consumption of oxygen and production of carbon dioxide) and photosynthesis (consumption of carbon dioxide and production of oxygen) can cause significant increases in diurnal dissolved oxygen (DO) and pH swings and result in decreased overall DO (Welch and Jacoby 2004, Anderson and Carpenter, 1998). Streams impacted by high levels of algal biomass will often demonstrate supersaturated DO concentrations and high pH values in late afternoon and minimum DO and pH values in early morning (Anderson and Carpenter, 1998). Low overnight DO concentrations can have considerable negative impacts on fish and in extreme cases the overnight low DO concentrations can be lethal for fish.

Adequate concentrations of dissolved oxygen are critical for the survival of fish. Decreased oxygen levels will increase the physiological stress of fish because their metabolic demands are not being met. This can impact growth and development at different life stages including eggs, alevins, and fry, as well as the swimming, feeding, and reproductive ability of juvenile and adult fish (Kramer 1987, Turner & Farley, 1971). Likewise, cool water fish, such as smallmouth bass, require clear streams with accessible cobble gravel substrate for feeding and spawning (Miller and Robinson, 2004).

The combination of increased nutrient loading and other factors, referred to as “cofactors”, together cause impacts (i.e. elevated algal growth, decreased DO, high pH), which lead to beneficial use impairments. The risk cofactors, in conjunction with nutrient loads, contribute to the degraded conditions manifested by the Illinois River watershed. Cofactors include light, temperature, flow, and canopy cover. Key cofactors in the Illinois River system are discussed below.

Riparian habitat serves several functions in stream systems including, providing shade and moderating water temperature. Riparian areas also serve to stabilize banks, prevent erosion, and add to overall stream channel complexity through inputs of woody debris and aid in pool formation (USDA, 2003). Reductions in riparian habitat have associated reductions in shade and increased water temperatures, which promotes the growth of algae and influences changes in DO and pH. Furthermore, channel alterations including erosion, straightening, and hardening prevent the river from maintaining productive stable stream banks and disconnect the river from riparian habitat thereby preventing an important riparian function - filtering runoff. Also, decreased flow conditions are more susceptible to high temperatures and low DO conditions.

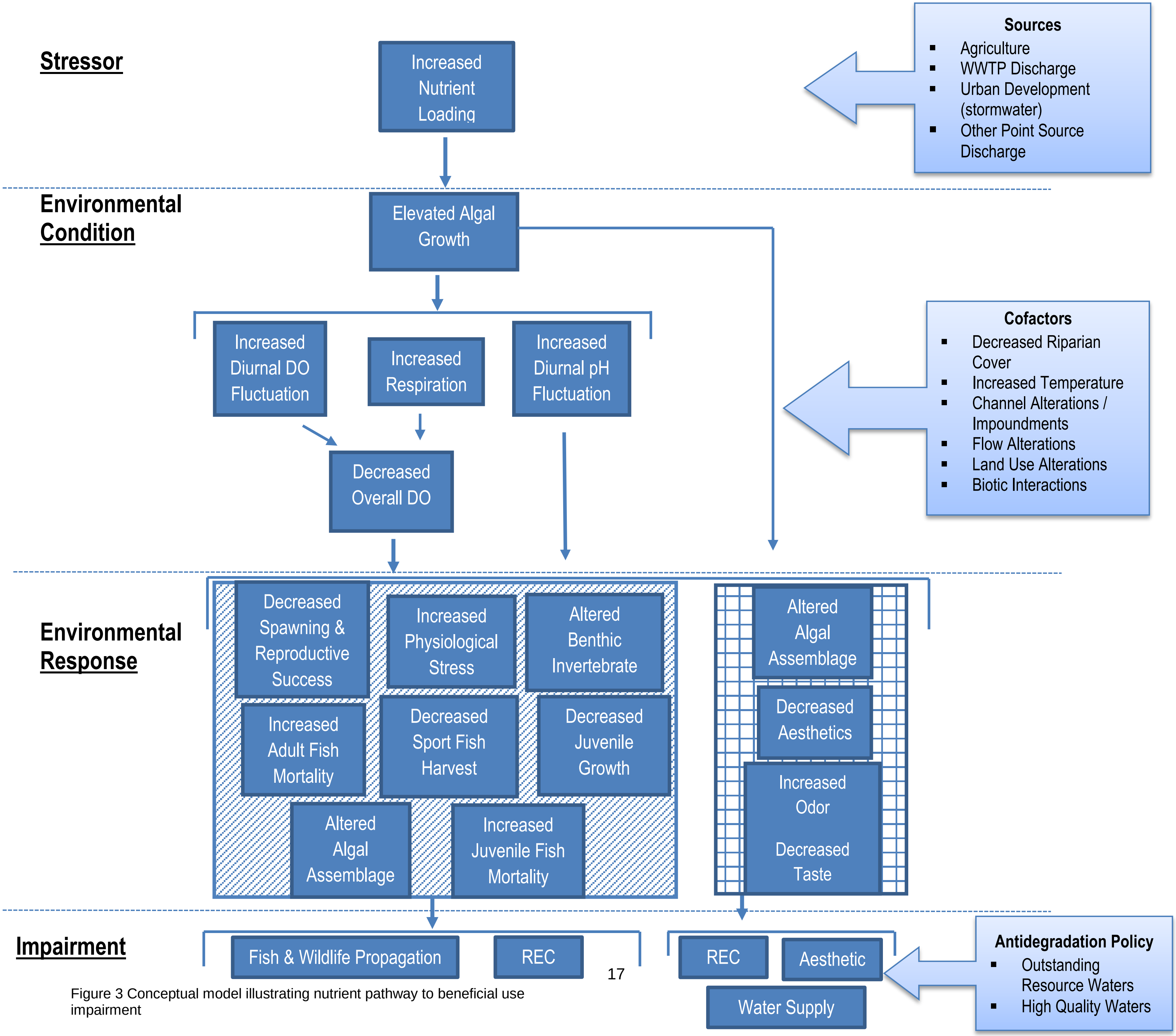


Figure 3 Conceptual model illustrating nutrient pathway to beneficial use impairment

Water Quality Standard Revision

Antidegradation Policy

Oklahoma’s WQS include protections for water quality through an Antidegradation Policy “*Waters of the state constitute a valuable resource and shall be protected, maintained, and improved for the benefit of all citizens. It is the policy of the State of Oklahoma to protect all waters of the state from degradation of water quality, as provided….”* (785:45-3-1). The WQS have afforded waterbodies broad protection under this policy since 1973 and in 1989 the antidegradation classes of Outstanding Resource Water (ORW), High Quality Water (HQW), and Sensitive Water Supply (SWS) were added. The Illinois River, Barren Fork Creek, and Flint Creek are all scenic rivers and classified as ORWs within Oklahoma’s Antidegradation Policy (785:45 Subchapter 3). Implementation of this policy requires that these waters are prohibited from receiving any new point source discharges or an increased load from any pollutant from an existing point source discharge (785:45-5-25(c)(1)).

The lower Illinois River (below Lake Tenkiller to the confluence with Arkansas River) and upper Illinois River, from Lake Tenkiller to confluence with Barren Fork Creek, are High Quality Waters (HQW) and also prohibited from any new point source discharges or an increased load from any pollutant from an existing point source discharge, unless it is demonstrated that the new discharge or increased load will result in maintaining or improving water quality (785:45-5-25 (3)). Additionally, no discharge of any pollutant to a HQW may lower existing water quality. To address nonpoint source discharges for both ORWs and HQWs the policy directs the implementation of management practices and conservation plans are required in subwatersheds where nonpoint source discharges are identified as causing or contributing to water quality degradation in an ORW. This WQS revision to the total phosphorus criterion in the Illinois River watershed does not change the Antidegradation Policy or its implementation.

Beneficial Uses

Beneficial uses establish the water quality goals for the waterbody. The beneficial uses for Illinois River, Barren Fork Creek, and Flint Creek are presented in Table 3. The total phosphorus criterion protects the Aesthetic beneficial uses in these waters. No changes to beneficial uses are made as part of this WQS revision.

Table 3 Designated beneficial uses for Illinois River, Barren Fork Creek, and Flint Creek

Waterbody Name	**Beneficial Uses**					
	Water Supply	**Fish & Wildlife Propagation**	**Agriculture**	**Recreation**	**Navigation**	**Aesthetics**
Illinois River	X	X	X	X		X
Barren Fork Creek	X	X	X	X		X
Flint Creek	X	X	X	X		X

Water Quality Criterion

Water quality criteria protect beneficial uses by setting limits on a pollutant (numeric criteria) or by describing an expected waterbody condition (narrative criteria); criteria have three components 1) magnitude, 2) duration, and 3) frequency. The numeric total phosphorus criterion applicable to Illinois River, Barren Fork Creek, and Flint Creek was established to protect the Aesthetic beneficial use and it is this criterion that is being revised. As presented in the Problem Identification section, nutrients and phosphorus specifically impact multiple beneficial uses; however, this total phosphorus criterion was originally adopted to specifically protect the Aesthetic beneficial use and this remains the case.

The Joint Study Committee made the following criterion recommendation in the Final Report (Joint Study, 2016).

> *"A six-month average total phosphorus level of not to exceed 0.035 mg/L"*

This recommendation addresses all three water quality criterion components.

- Magnitude: 0.035 mg/L
- Duration: six-month average
- Frequency: not to exceed (zero exceedance allowance)

Each of the criterion components are presented below along with OWRB staff analysis and recommendation.

Criterion Magnitude

The criterion magnitude defines the amount of a pollutant that can be allowed while still maintaining the waterbody's beneficial uses. The total phosphorus criterion magnitude is 0.037 mg/L and this magnitude was not revised. The original adoption and subsequent review of this value was scientifically defensible and protective of the beneficial use and was also approved by U. S. EPA Region 6. The results of the Joint Study provide further evidence that this value is ecologically relevant; for example, Figure 33 in the 2016 Joint Study final report presents the greatest number of total phosphorus (as a 6-month average) change points associated with various ecological response variables at the concentration of 0.037 mg/L.

In the *Second Statement of Joint Principles and Actions*, it was agreed that OWRB would only be required to revise the total phosphorus criterion if the committee recommended a value that was "significantly different" from the value of 0.037 mg/L. In this context the term "significantly different" does not have any statistical connotation and only refers to the ± 0.01 interval set forth in the *Second Statement of Joint Principles and Actions.* Meaning if the committee recommended a total phosphorus concentration criterion magnitude at or between 0.027 mg/L and 0.047 mg/L OWRB was not required to revise the criterion. The committee recommended criterion magnitude was 0.035 mg/L; therefore, OWRB elected not to revise the criterion magnitude.

Criterion Duration

The criterion duration specifies the maximum allowable time period over which receiving water concentrations can be averaged for comparison with the magnitude value (i.e. an averaging period). The current criterion has a 30-day geometric mean averaging period and staff at that time found this period to be suitable (OWRB, 2012). However, in the water quality assessment program, since approximately 2007 the criterion duration has been implemented as a 90-day geometric mean (785:46-15-14). The decision to employ a 90-day geometric mean averaging period was made to accommodate monitoring programs and allow a greater time window for sampling events to take place. The ecological relevance of these averaging periods were never evaluated. Moreover, the current criterion has not been implemented in other regulatory programs such as permitting.

In the Joint Study, duration was considered by evaluating the relationship between mean total phosphorus calculated at different monthly periods and the algal biomass response variable. The monthly periods for average TP concentration included 2, 4, 6, 8, 10, and 12-month periods. The TP change point was calculated for each averaging period. Change point analysis is used to identify a threshold in a relationship between two variables; in this case, the variables were TP concentration and algal biomass (EPA, 2020). The analysis works by finding a point along a distribution (here TP concentrations) where characteristics of values before and after this point are different. In this case, the analysis identified a TP concentration where the algal biomass concentration was considerably different before and after that TP concentration point. The Interim Report from April 2016 presents the 2 through 12-month TP averaging periods and it can be observed that the variability in the TP change points decreases as the averaging period increases. The six-month averaging period was seen to be relatively similar to the longer periods (8, 10, & 12 months) and a more consistent metric than shorter averaging periods (2 & 4 months). Thus the 6-month average was selected as the criterion duration because this period for average TP concentration was identified as a consistent predictor of shifts in algal biomass concentration.

Additionally, the TP concentration and algal biomass deviations from the median two-year values of each parameter were evaluated. This analysis demonstrated that when the algal biomass value was greater than the algal biomass median, the total phosphorus concentration was often considerably lower than the total phosphorus median concentration. Similarly, it was observed that the majority of high algal biomass measurements corresponded to reductions in TP concentration (Joint Study 2016). It is clear from these two analyses that antecedent TP concentrations support blooms of algal biomass. This supports the use of the six-month averaging period because it provides for a time integrated evaluation of TP concentrations (stressor) driving the algal biomass response, which impacts the beneficial use.

Statistical Analyses

OWRB staff conducted additional analyses to gauge the application of a six-month average in the state’s water quality assessment program. Exploratory data analysis and hypothesis testing were used to evaluate different averaging periods. USGS and OWRB conduct regular water quality monitoring (monthly sampling since 1999) in the Illinois River watershed and on other scenic rivers. This monthly data was used for these analyses; often a data period of 2008-2018 was used and for some analysis a period of 1999-2018 was used. As part of the current monitoring program, generally six high flow targeted monitoring events occur annually. This analysis excluded total phosphorus data from the target high flow monitoring events.

Table 4 Water quality monitoring stations used in criterion duration statistical analyses

Station Name	Station ID	Data Providers	Period of Record Used
Illinois River near Watts	7195500	OWRB/USGS	1999-2018; 2008-2018
Illinois River near Tahlequah	7196500	OWRB/USGS	1999-2018; 2008-2018
Flint Creek near Kansas	7196000	OWRB/USGS	1999-2018; 2008-2018
Barren Fork Creek near Eldon	7197000	OWRB/USGS	1999-2018; 2008-2018
Lee Creek near Short	7249985	OWRB	2000-2018
Mountain Fork near Smithville	7338750	OWRB	2000-2018

Exploratory data analyses were used to determine if: 1) there were notable differences in averaging periods, and 2) if these differences were consistent across the waterbodies. Initially, the probability of exceeding the criterion magnitude of 0.037 mg/L at various averaging periods (3, 4, 6, and 12 months) was considered (Figure 4). Generally, the averaging periods have different exceedance probabilities. However, the general magnitude of the exceedance probabilities and the effect of the averaging period were related to the underlying dataset. For example, for the Illinois River Stations (both Watts and Tahlequah), the greater majority of ambient data are above the 0.037 mg/L criterion magnitude, which increases the magnitude of the exceedance probabilities, regardless of the averaging period. Furthermore, the probability of an exceedance increases as the averaging period increases from 3 to 12 months. However, an opposite pattern exists for Barren Fork Creek, where the greater majority of ambient data are below 0.037 mg/L. Expectedly, the magnitudes of the exceedance probabilities are considerably less. However, as the averaging period increases from 3 months to 12 months, the likelihood of exceeding the criterion decreases.

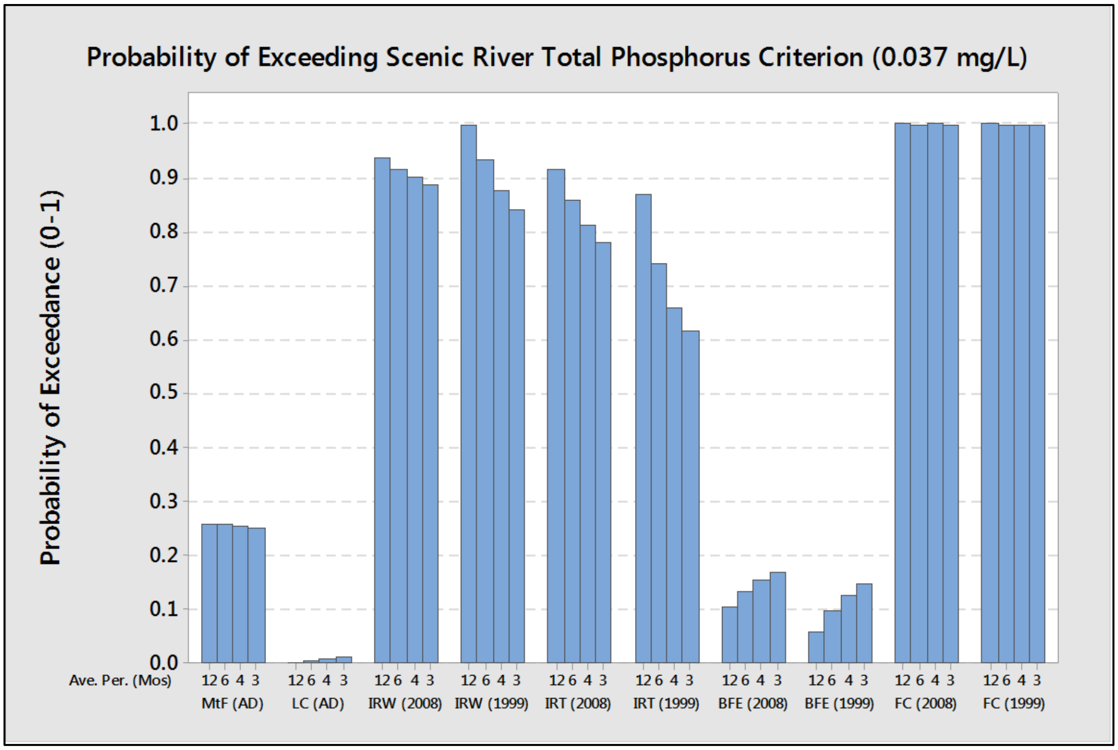


Figure 4 Probability of exceedance of the 37 ug/L at 3, 4, 6, and 12 month averaging periods for Oklahoma's scenic rivers. For the Illinois River (IRW and IRT), Barren Fork Creek (BFE), and Flint Creek (FC), data periods from 1999-2018 and 2008-2019 were used.

To better understand these differences, additional exploratory analysis compared the distribution of averaged data within the various averaging periods for both the Illinois River near Watts and Barren Fork Creek. These two stations represent the opposite distributional effects discussed in the preceding paragraph. For each station, a similar period of record was used (2008-2018). For each dataset, rolling means were calculated using 2, 3, 4, 5, 6, 9, and 12-month periods. Using the calculated rolling means, distributions for each averaging period at each station were created using boxplots (Figures 5 and 6). The boxplots were constructed using the 10th and 90th percentiles as the lower and upper boundary of the box, and the 1st and 99th percentiles for the corresponding whiskers. Dataset means and medians are represented by the darkened circle and the crosshair circle, respectively.

Initial observations of the data showed a consistent pattern for the 2 stations. As the averaging period increases, data become more normally distributed (mean and median become more similar) and have a narrower distribution. However, when comparing to the water quality criterion magnitude of 0.037 mg/L (red dotted line in each figure), averaging periods may introduce decision bias. This is conveniently illustrated using simple hypothesis testing and the associated likelihood of a Type I (rejecting the null hypothesis when true) error or a Type II (not rejecting the null hypothesis when false) error. For water quality implementation programs, the null hypothesis would be the total phosphorus data are meeting the water quality criterion using a given averaging period. When looking for water quality impairments, an associated Type I or alpha error would determine the waterbody is impaired when it is, in fact, not impaired. Conversely, a Type II or beta error would determine that the waterbody is not impaired when, in fact, it is impaired. For water quality implementation programs, it is imperative that Type II errors are avoided so that protection is not decreased. Using the Illinois River and the Barren Fork Creek (Figures 5 and 6), two scenarios illustrate the opposite effects of Type II decision bias.

- Scenario 1: For the Illinois River near Watts (Figure 5), most ambient data are above the 0.037 mg/L ug/L criterion magnitude. As averaging periods decrease (from 12 to 2 months), the likelihood of a Type II decision error increases. Protection increases as averaging periods increase.

- Scenario 2: For Barren Fork Creek (Figure 6), most ambient data are below the 0.037 mg/L ug/L criterion magnitude. As averaging periods increase (from 2 to 12 months), the likelihood of a Type II decision error increases. Protection increases as averaging periods decrease.

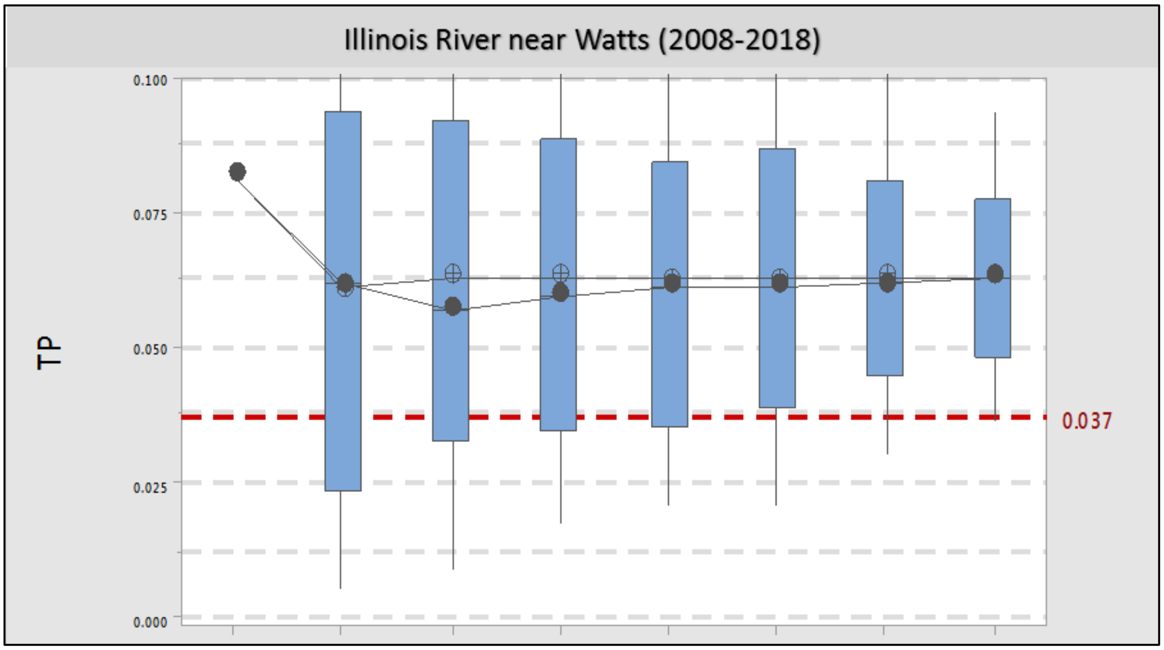


Figure 5 Distributional analysis of the Illinois River near Watts (2008-2018). Data represent averaging periods, including (in boxplot bars from left to right): 2, 3, 4, 5, 6, 9, and 12 months. Each boxplot represents the distribution of the various averages at a particular averaging period. Each box represents the 10th and 90th percentile. Dataset means and medians are represented by the darkened circle and the crosshair circle.

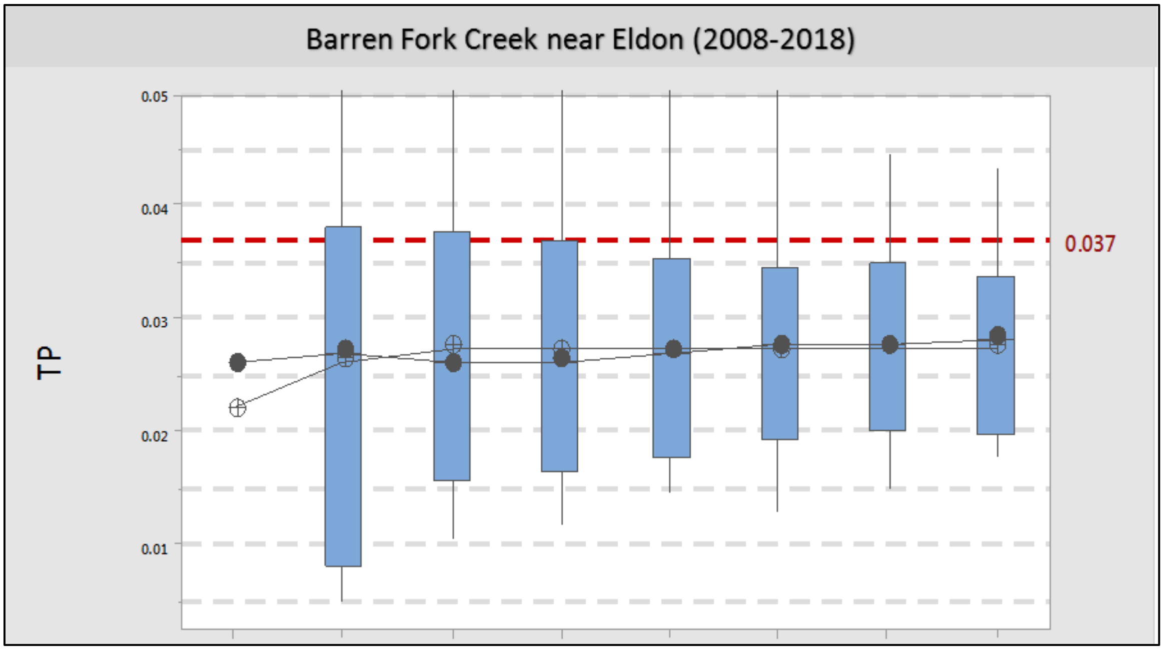


Figure 6 Distributional analysis of the Barren Fork Creek near Eldon (2008-2018). Data represent averaging periods, including (in boxplot bars from left to right): 2, 3, 4, 5, 6, 9, and 12 months. Each boxplot represents the distribution of the various averages at a particular averaging period. Each box represents the 10th and 90th percentile. Dataset means and medians are represented by the darkened circle and the crosshair circle.

To this point, exploratory analysis has elucidated several important factors when considering averaging periods. First, decision bias is affected by the proximity of both the raw and the averaged data to the criterion. Depending on the relative proximity of data to the criterion, averaging periods at both ends of the spectrum (2 or 12 months) can decrease protection. This was demonstrated in the preceding analysis and needs no further exploration. Second, averaging periods have notable differences when considering the magnitude of exceedance probability, which leads to several unanswered questions. Even though visual differences occur between averaging periods, was there an actual significant difference between averaging periods and what is the magnitude of that difference? To answer these questions, different hypothesis testing approaches were used.

To test the differences between means calculated at various averaging periods, a fixed effects analysis of variance (AVOVA) was used. Water quality data generally do not meet the underlying assumption of data normality necessary to use a parametric test, such as an ANOVA. However, when using averaged data as a test dataset, normality increases, especially as the averaging periods increase (Figures 5 and 6 above). At both monitoring stations, means and medians become nearly equivalent and the distribution becomes increasingly mesokurtic moving from a 2 month to 12 month averaging period. The ANOVA was employed as a general linear model (GLM) using both averaging period and season as categorical predictors, as well as a combined effects predictor (averaging period*season). The model produced least squared (fitted) means for analysis. Using standard error of the least squared means, each predictor was tested for significance (Figure 7 and 8).

To simplify discussion, Illinois River at Tahlequah and Barren Fork Creek are displayed and used for analysis. Both categorical predictors (season and averaging period) showed statistically significant differences across various categorical predictors (Figure 7). For averaging periods, a general grouping of shorter averaging periods (2 to 4 months) were significantly different than longer averaging periods (5 to 12 months). For season, summer and spring seasons showed significant differences, while the fall/winter season showed inconsistency in grouping. Since both season and averaging period were statistically significant categorical predictors, the combined effects predictor (averaging period*season) was also used for analysis (Figure 8). Only spring and summer are analyzed as combined effects predictors. The averaging period groupings are relatively inconsistent across between stations and seasons. However, in all instances, a consistent transition point does exist between the 5 and 6 month averaging periods. Furthermore, although significant differences do exist, the relative magnitudes of the differences were small in comparison to the 0.037 mg/L criterion. To explore this relative magnitude of difference, the simplest procedure was to compare the LS means, as well as standard error bounds on the lowest LS mean and the highest LS mean (Table 5). Generally, these differences were relatively low. Additionally, when looking at the recommended 6 month duration, its relative magnitude generally lies between the lowest and highest LS mean. Although statistically significant, the magnitude of differences between the individual averaging periods was generally small.

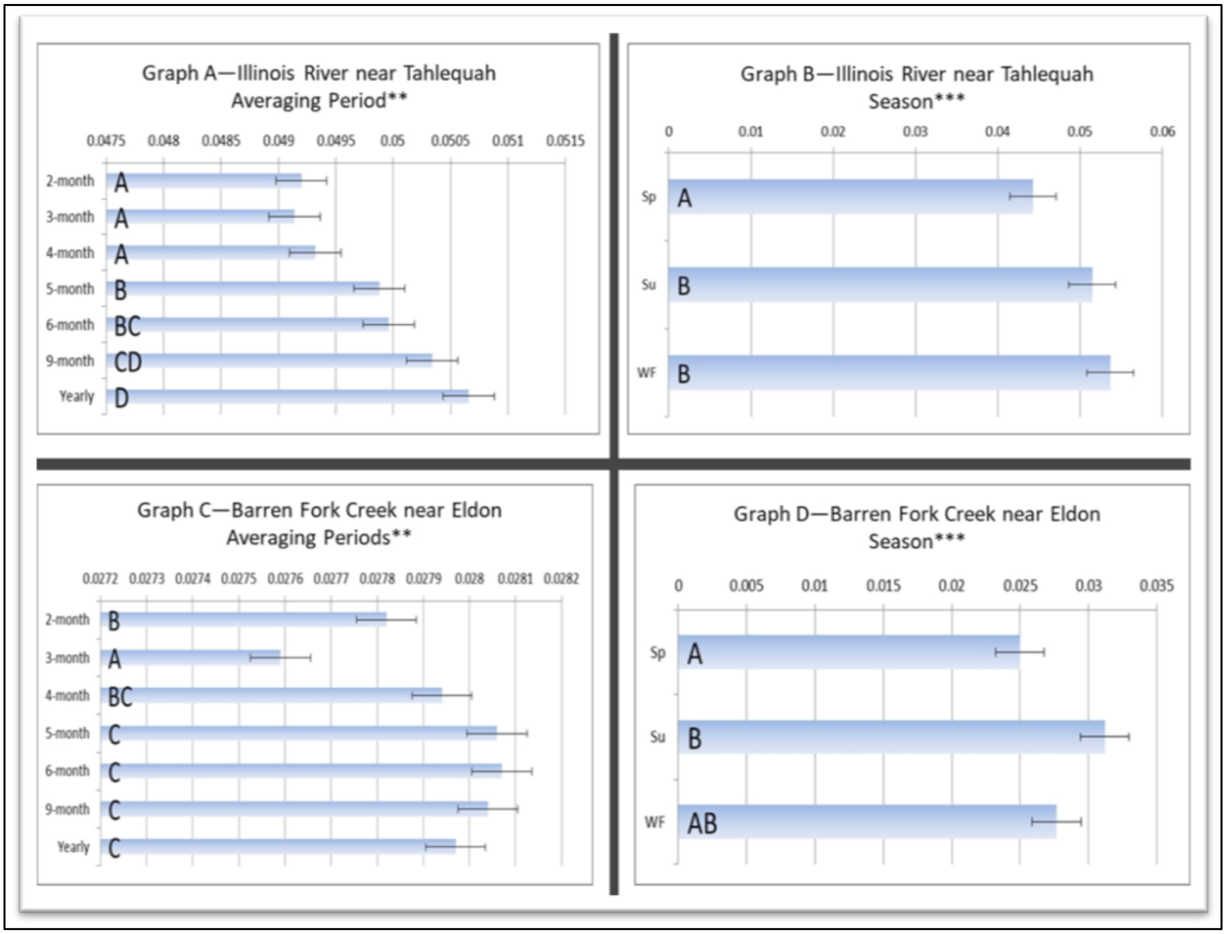


Figure 7 Analysis of Variance of data means at various averaging periods and seasons for the Illinois River near Tahlequah (2008-2018) and Barren Fork Creek near Eldon (2008-2018). Graphs A and C show LS means for averaging period categorical predictor. Graphs B and D show LS means for seasonal categorical predictor. Error bars represent standard error. Letters on the left margin of each graph show where both overlap and difference in statistical significance occurs.

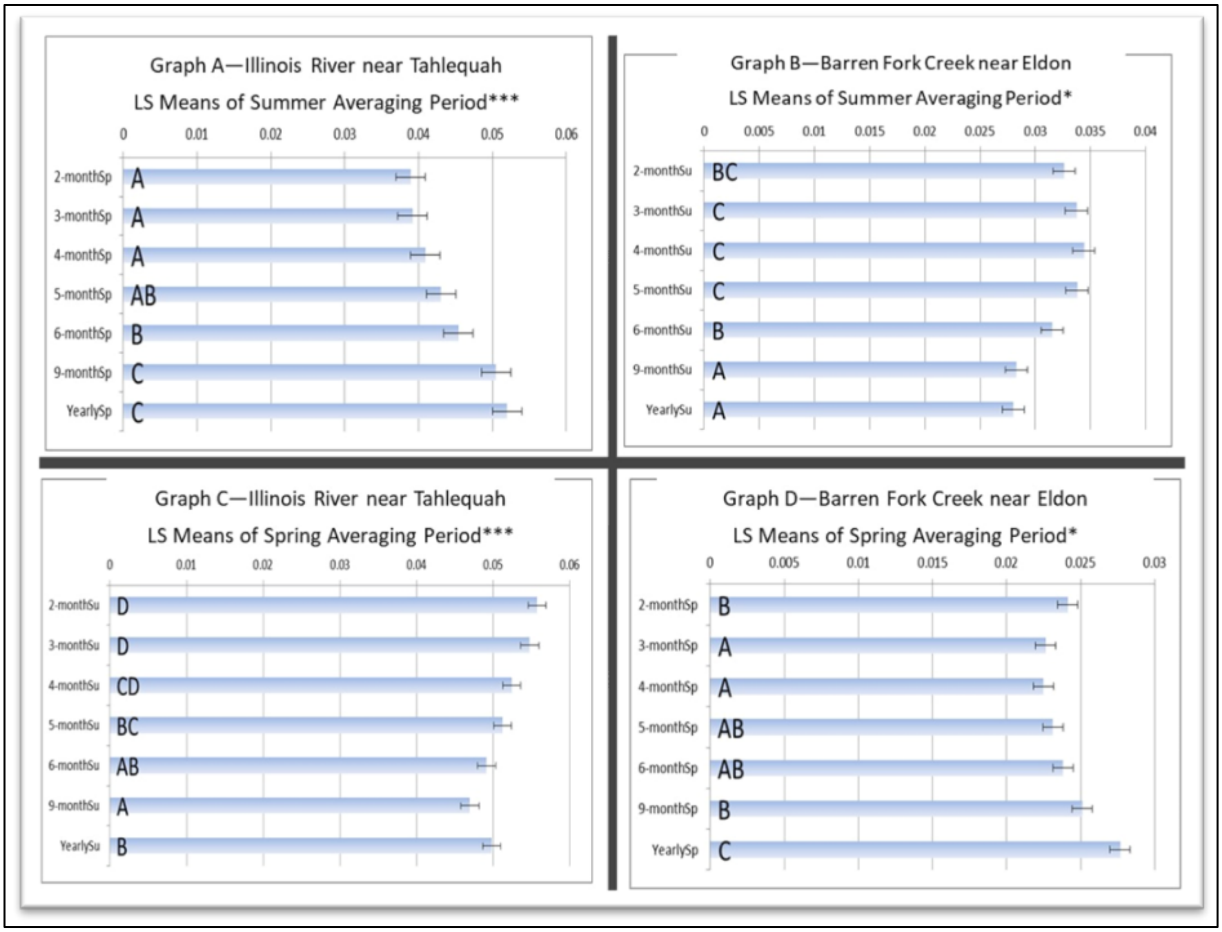


Figure 8 Analysis of Variance of data means for combined effects predictor (averaging period*season) for the Illinois River near Tahlequah (2008-2018) and Barren Fork Creek near Eldon (2008-2018). Graphs A and B show LS means for the summer combined predictor at both stations. Graphs C and D show LS means for the spring combined predictor at both stations. Error bars represent standard error. Letters on the left margin of each graph show where both overlap and difference in statistical significance occurs.

Table 5 Difference between smallest and largest LS Means in Figure 5

Station	Season	Phosphorus LS Mean Difference (ug/L)	Phosphorus SE Bounds Difference (ug/L)
Illinois River near Tahlequah	Summer	8	5
Illinois River near Tahlequah	Spring	12	9
Barren Fork Creek	Summer	9	6
Barren Fork Creek	Spring	5	3

For 303(d) assessment implementation, some nth percentile of data is typically used to determine compliance. For instance, Oklahoma's Use Support Assessment Protocols (USAP) (OAC 785:46-15) typically uses the 10th percentile of data for assessment. If greater than 10 percent of data are above a criterion, the beneficial use is impaired. For this reason, use of an ANOVA becomes impractical for comparative purposes. However, a similar hypothesis analysis can be used for distributions. For any empirical probability curve, a confidence interval may be calculated around the curve. Using these confidence intervals allows for comparison of averaging periods at any point in the distribution. For assessment of the total phosphorus criterion, the 10th and the 90th percentiles of the distribution can be used for analysis. The lower bound, or the 10th percentile can be representative of stations like the Illinois River at Tahlequah. In this instance, the differences in the 10th percentile at various averaging periods becomes important because the preponderance of data lie above the 0.037 mg/L criterion, so the potential bias exists at the lower tail of the distribution. Conversely, the upper bound, or the 90th percentile, can be representative of stations like Barren Fork Creek. In this instance, the differences in the 90th percentile at various averaging periods becomes important because the preponderance of data lie below the 0.037 mg/L criterion, so the potential bias exists at the upper tail of the distribution. For comparison purposes, both tails of the distribution for both stations are shown in Figure 9. The averaging periods are significantly different for both stations at the 10th and the 90th percentiles. And unlike the analysis of LS means, the differences are quite notable. However, in all instances, a consistent point of transition between shorter and longer averaging periods lies at the 5 to 6 month averaging period and the 5 to 6 month averaging shows significant grouping to both tails of the distribution.

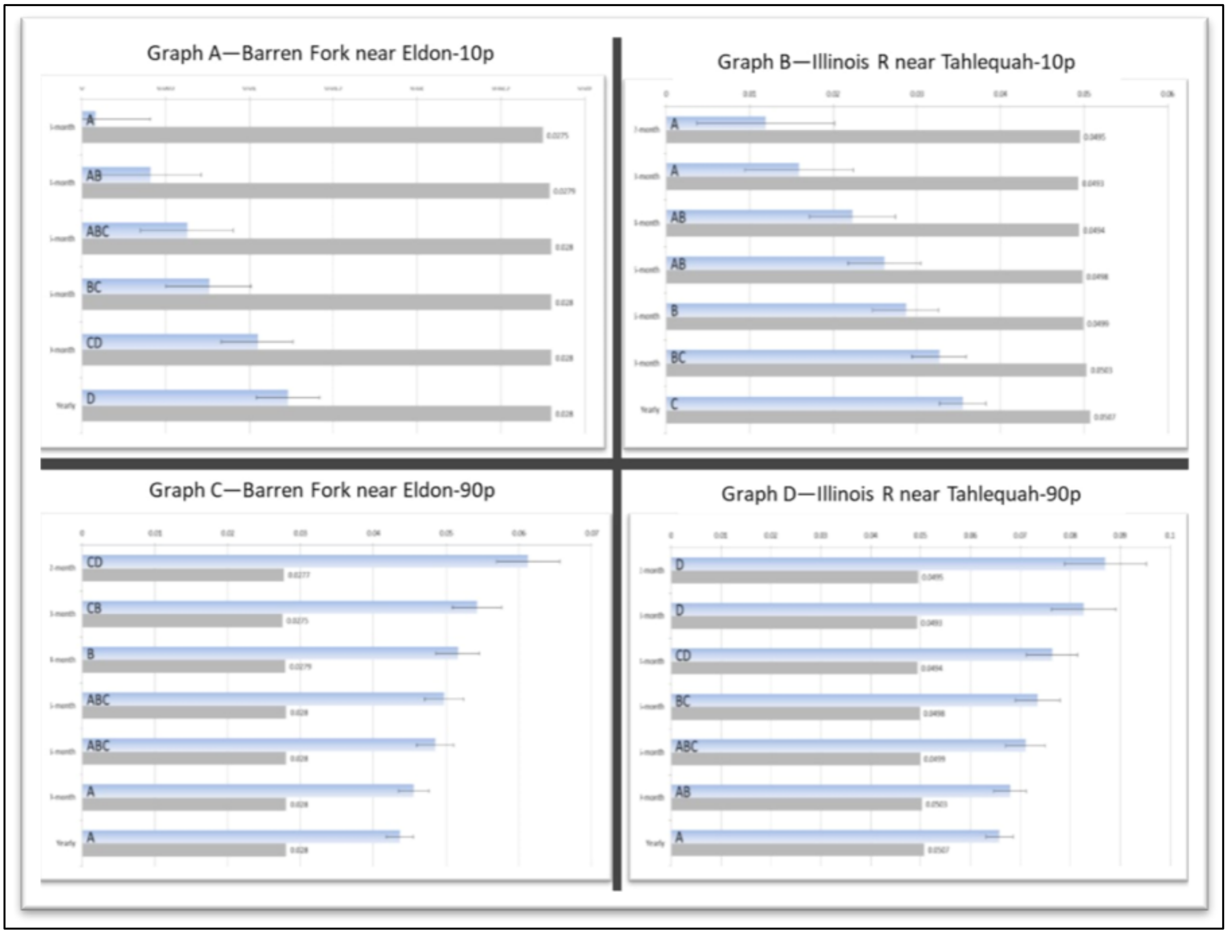


Figure 9 Analysis of data 10th and 90th percentiles at various averaging periods for the Illinois River near Tahlequah (2008-2018) and Barren Fork Creek near Eldon (2008-2018). Graphics show the data mean in the gray shaded bar and the data percentile for total phosphorus in the blue bar. Graphs A and B represent data at the 10th percentile. Graphs C and D represent graphs at the 90th percentile. Error bars represent the 95% confidence interval around the distribution. Letters on the left margin of each graph show where both overlap and difference in statistical significance occurs.

Conclusion

The data analysis described above supports a 6-month average as the duration component of the criterion. This averaging period protects the Aesthetics beneficial use and has practical application for use in water quality management implementation programs. First, the 6-month average period effectively integrates stress inter-seasonally, across periods of both phosphorus loading and biological uptake. However, it is not unnecessarily long in duration, allowing for higher phosphorus values to be muted in the overall average. Second, when looking at various implementation approaches, some significant differences do exist across the various averaging periods. When comparing means, the observed differences are small; yet, in the comparison of percentiles, the differences do become larger. However, in both instances, the 6-month averaging period typically provides a notable transition between lower and higher averaging periods. Although significance does occur above and below this transition, the relative magnitude of the differences becomes much smaller. Furthermore, the middle averaging periods (4, 5, 6, months) sometimes show grouping with one or both tails of the distribution. Third, analyses demonstrate that decision bias occurs when using both shorter (2 to 3 month) and longer (9 to 12 month) averaging periods and was dependent on whether data are generally above or below the 0.037 mg/L criterion magnitude. The 6-month average represents a middle of the road averaging period to integrate both data situations. Although some decision-making bias may still occur with a 6-month averaging period, it is significantly lessened at this averaging point.

Criterion Frequency

In ecology the term resilience is broadly defined as the ability of a system to recover following stress or disturbance. The rate of system recovery is dependent on the severity of the disturbance and its extent. Generally, phosphorus criteria are set at concentrations within a range of natural variability and a short term spike in in-stream phosphorus concentration does not inevitably impair the beneficial use. It is typical for stream systems to experience variable phosphorus inputs related to various flow conditions (i.e., wet weather vs. dry weather). Although there can be a dramatic short term response (i.e., algal bloom) to a spike in phosphorus concentration, a single or a few increased phosphorus events do not cause long term shifts in biological conditions, such as dominance by algal taxa associated with increased phosphorus. For example, the algal species *Cladophora glomerata* is indicative of streams with increased ambient phosphorus concentrations and *Spirogyra sp.* are indicative of streams with low phosphorus concentrations (Fetscher, 2014). Moreover, the results of the TITAN analysis in the Joint Study (2016) reported that community threshold for negative responding algal taxa (taxa that increase with increasing TP concentration) was 0.021 mg/L TP.

Algal community dominance by *Cladophora glomerata* has been documented in the Illinois River watershed. In the Illinois River mainstem, benthic algal biomass values were typically about 350 mg/m^2 and at times values were as high as 2,000 mg/m^2 (Joint Study, 2016). Additionally, it was found that at algal biomass values of approximately 180-300 mg/m^2, the biovolume of *Cladophora glomerata* dramatically increased (Joint Study, 2016). The serious phosphorus impacts and impairment of beneficial uses (see figure 3) is the result of ongoing exposure to elevated phosphorus concentrations.

The criterion component of frequency is the allowable number of excursions from the magnitude and duration that can occur within a particular time period and still protect the beneficial use. The frequency for both the current criterion and the recommendation from the Joint Study Committee are "*never to exceed*" meaning that there are zero allowable excursions from the magnitude of 0.037 mg/L and the duration of 6-month average. This implies that any excursion would cause an impairment of the waterbody's beneficial use, which does not align with the principle of ecological resilience and short term versus ongoing exposure to phosphorus pollution. Moreover, the *never to exceed* approach to frequency does not provide any flexibility to implementation programs and overly rigid criteria are often not effectively implemented. Therefore, excursions will be allowed in a two-part approach to criterion frequency that also serves to protect the beneficial use. The criterion magnitude and duration shall not be exceeded more than once in a one-year period (part 1) and not more than three times in a five-year period (part 2). This frequency recognized the ability of ecosystems to process short term increased nutrient concentrations while guarding against the cumulative chronic effects that impair beneficial uses.

Although the original criterion was adopted with a frequency of *never to exceed*, this frequency has never been implemented. In the water quality assessment program, since approximately 2007, the criterion frequency has been implemented as a 25% exceedance allowance (785:46-15-14). Table 6 below presents an evaluation of beneficial use attainment according to the different frequencies associated with the criterion. At this time, change to the criterion frequency has very little impact on the assessment outcome of whether or not the beneficial use is in attainment because the waterbodies are consistently subject to high amounts of phosphorus pollution. Over the years, marked reductions in phosphorus have been observed in Barren Fork Creek and therefore, determination of beneficial use attainment can be variable based on the underlying data set. Lee Creek is another Oklahoma scenic river generally regarded to be in very good condition and was included here for comparative context. Under the *never to exceed* approach even Lee Creek was identified as impaired based on 2 samples over the 5-year span.

Table 6 Waterbody beneficial use attainment according to various approaches for criterion frequency

Monitoring Location	**Frequency Approach** (Five Year Assessment Period (2013-2018))		
	Never to Exceed	**25% Exceedance Allowance**	**Current Recommendation**
Illinois River at Watts	Ben. Use Impaired	Ben. Use Impaired	Ben. Use Impaired
Illinois River at Tahlequah	Ben. Use Impaired	Ben. Use Impaired	Ben. Use Impaired
Barren Fork Creek	Ben. Use Impaired	Variable Outcome	Ben. Use Impaired
Lee Creek*	Ben. Use Impaired	Ben. Use Attaining	Ben. Use Attaining

*Lee Creek an Oklahoma scenic river included here for comparative context.

Final Revised Criterion

Combining all three components (magnitude, duration & frequency) the final revised criterion applicable to Illinois River, Barren Fork Creek, and Flint Creek is:

> *The total phosphorus six month rolling average of 0.037 mg/L shall not be exceeded more than once in a one-year period and not more than three times in a five-year period.*

Information presented above supports the staff finding that this criterion will protect the Aesthetic beneficial use of Illinois River, Barren Fork Creek, and Flint Creek. This language is recommended for adoption into the Oklahoma Water Quality Standards in section 785:45-5-19.

Critical Condition

As described in the Environmental Setting and Background section, the Joint Study Committee made the following recommendation in the Final Report (Joint Study, 2016).

> *"A six-month average total phosphorus level of not to exceed 0.035 mg/L based on water samples taken during the critical condition, as previously defined…"*

The Joint Study Committee defined the term critical condition as *“the conditions where surface runoff is not the dominant influence of total flow and stream ecosystem processes.”* The critical condition term introduced by the Joint Study Committee is a new term for Oklahoma Water Quality Standards and has a narrow application to WQS implementation. The term instructs which water quality sample results should be utilized to evaluate the total phosphorus criterion for the purposes of water quality assessment. A technical analysis was needed to translate the committee critical condition terminology into an operational definition that could be feasibly and consistently implemented by water quality management programs across multiple agencies in both states. Moreover, the technical analysis evaluated the impact of the new critical condition term on how total phosphorus water quality conditions would be characterized within the waterbody.

The Joint Study Committee critical condition definition speaks to two endpoints 1) total flow and 2) stream ecosystem processes; therefore, independent analyses were conducted to address each endpoint. Hydrograph separation analysis was used to evaluate the total flow endpoint and a scour analysis was used to evaluate the stream ecosystem process endpoint. These analyses were conducted by OWRB staff, in consultation with Oklahoma sister environmental agencies[1], the Cherokee Nation, and Arkansas Division of Environmental Quality.

Streamflow Characterization

A hydrograph is a graph of streamflow or discharge over time, typically with the units of cubic feet per second (cfs). Figure 10 presents a hydrograph for the Illinois River at Tahlequah. Streamflow is composed of a combination of baseflow (return flow from groundwater), interflow (rapid subsurface flow), and overland flow (surface flow over poorly permeable or temporarily saturated soils). Together interflow and overland flow are known as quickflow or direct runoff, which is the rapid runoff of “new” water into the stream channel during a rain event. Baseflow typically reaches the stream through a longer flow path and sustains streamflow during periods without rain. Baseflow can be dynamic and influenced by seasonal factors.

---

[1] Sister environmental agencies includes the OK Department of Environmental Quality, OK Conservation Commission, OK Department of Agriculture Food & Forestry, OK Department of Wildlife Conservation

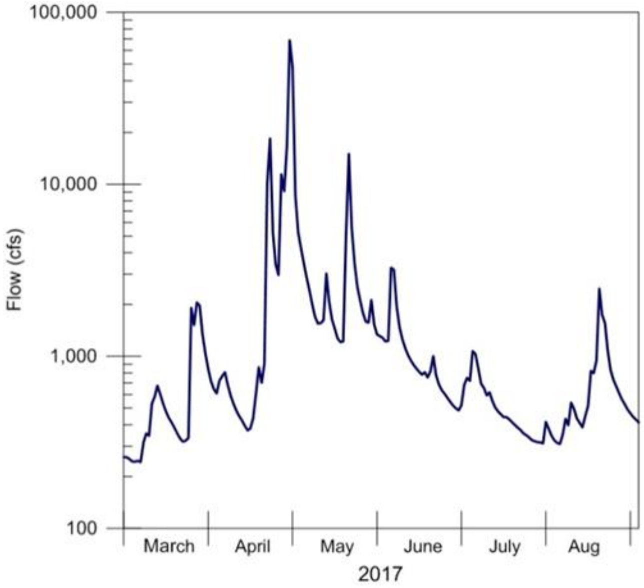


Figure 10 Hydrograph: Illinois River at Tahlequah March - August 2017

Hydrograph Separation Methods

Hydrograph separation is a procedure that partitions the hydrograph into two key component flows, baseflow and direct runoff. Figure 11 is again the hydrograph for the Illinois River at Tahlequah, but the total hydrograph has been separated to show the portion that is baseflow.

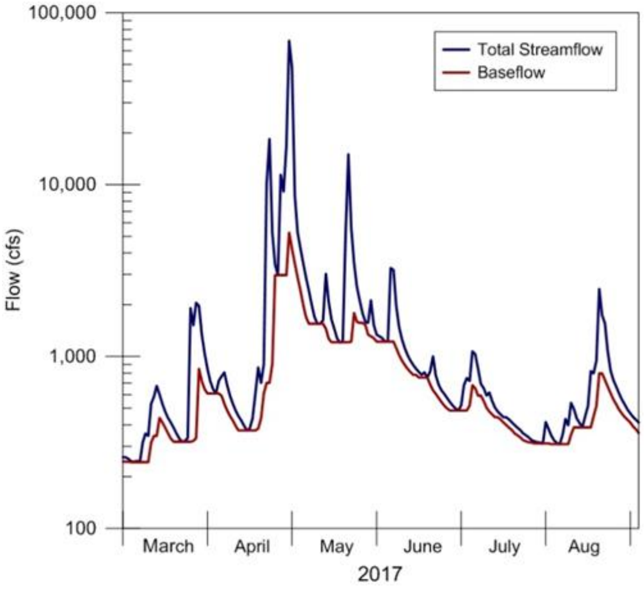


Figure 11 Illinois River at Tahlequah, hydrograph of total streamflow and baseflow (March- August 2017); baseflow according to HYSEP Sliding-Interval method

Hydrograph separation is a tool that has long been used by Hydrologists in an effort to identify and understand the components of streamflow and runoff generation processes (i.e., source areas, pathways, retention times) (Gonzales et al., 2009). An estimation of baseflow versus direct runoff is useful to understand overall watershed hydrology, such as flood and low flow conditions and groundwater surface water interactions. Hydrograph separation methods can be divided into two major approaches 1) tracer-based methods and 2) non-tracer-based methods. Tracer-based methods employ various chemicophysical signatures of water and physical processes to differentiate the contributions of baseflow and direct runoff to total flow. Tracer-based methods will not be discussed further in this document. Non-tracer-based methods include graphical hydrograph separation methods. These methods use the hydrograph itself as a signal and utilize time as a function with the assumption that the time of a direct runoff event is much shorter than that of groundwater discharge and this time is relatively constant across rain events (Pelletier, 2019). Three USGS graphical hydrograph separation methods and one alternative method (McCarty and Haggard, 2016), termed here as the Delta 10% method were considered as part of this project.

- HySEP Local Minimum
- HySEP Sliding-interval
- PART
- Delta 10%

These methods separate a streamflow hydrograph based on a mathematical technique and the USGS Groundwater Toolbox (Barlow et al., 2015) computer program was used to ensure consistency and efficiency. The Delta 10% method was conducted in Microsoft Excel. However, it is important to note that even though the methods are formal algorithms for identifying baseflow versus total streamflow the methods are subjective and not based on physical processes as used in tracer-based methods. Listed below are several underlying assumptions applicable to these methods.

- Flow systems are driven by diffuse areal recharge uniformly disturbed over a watershed
- Single point of outflow from the basin at the gauging station of interest
- All groundwater in the basin discharges to the stream, except that lost by evapotranspiration
- Streamflow hydrograph represents water contributions from two sources: surface runoff and groundwater discharge from a single aquifer

- Groundwater and surface water drainage areas are coincident
- Regulation or diversion of flow should be minimal and groundwater pumping minimal
- Several years of record should be analyzed

The three USGS methods and the Delta 10% method are described below. A calculation of duration of surface runoff is foundational to all three USGS methods. The duration of surface runoff is calculated using the following empirical relationship.

$$N=A^{0.2}$$

N is the number of days after which surface runoff ends and A is the watershed area in square miles (Lindsley et al., 1975 and Sloto and Michele, 1996).

HYSEP Sliding-interval

The USGS HYSEP hydrograph separation methods utilize an interval of days (2N*) defined as the odd integer between 3 and 11 nearest to 2N. In the sliding-interval method, baseflow is assigned as the lowest daily discharge that occurs within the interval [0.5(2N*-1) days] before the day of interest and after the day of interest. Using Barren Fork Creek as an example, if the day of interest is May 4, 2017, the method looks back 3 days and forward 3 days to identify the lowest daily discharge and assigns that discharge to May 4th as the baseflow (Table 7). In this example the baseflow value of 511 cfs was assigned to May 4, 2017. Figure 12 shows the hydrograph of total flow and baseflow for May 2017 using the HYSEP sliding-interval method.

Table 7 Example determining baseflow value according to HYSEP Sliding-interval method

Day of Interest: May 4, 2017	
Watershed Area: 307 square miles	
Interval: 3 days	
Date	Total Streamflow (cfs)
May 1, 2017	2430
May 2, 2017	1690
May 3, 2017	1190
May 4, 2017	894
May 5, 2017	725
May 6, 2017	601
May 7, 2017	511
Baseflow assigned to May 4, 2017	511

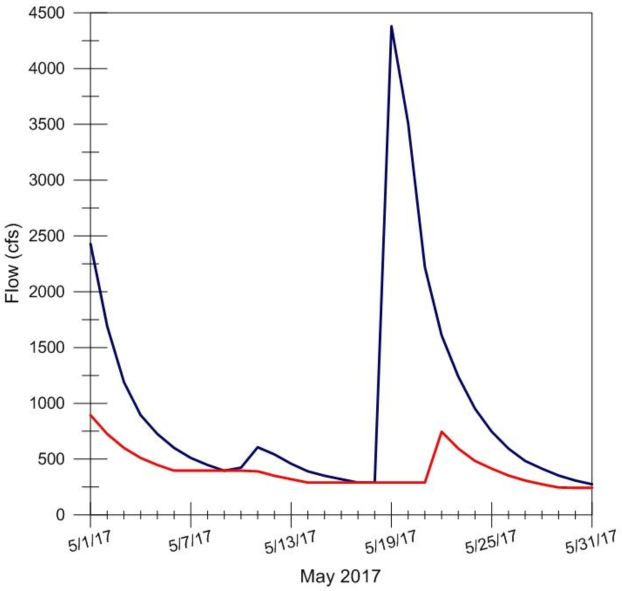


Figure 12 Barren Fork at Eldon, separated hydrograph (streamflow and baseflow) according to HYSEP Sliding-interval method

HYSEP Local minimum

The local minimum method evaluates each day to determine if it has the lowest discharge in the interval [0.5(2N*-1) days] before and after the given day. If it is the day with the lowest discharge in the total interval, then the day is a local minimum. Adjacent local minimums are connected with straight lines and baseflow values for days in between local minimums are estimated by interpolation (Table 8, and Figure 13).

Table 8 Example determining baseflow value according to HYSEP Local minimum method

Watershed Area: 307 square miles	
Interval: 3 days	
Date	Total Streamflow (cfs)
May 6, 2017	601
May 7, 2017	511
May 8, 2017	448
May 9, 2017	397 (local minimum)
May 10, 2017	424
May 11, 2017	607
May 12, 2017	542
May 13, 2017	459
May 14, 2017	391
May 15, 2017	351
May 16,2017	320
May 17, 2017	290 (local minimum)
May 18, 2017	290
May 19, 2017	4380

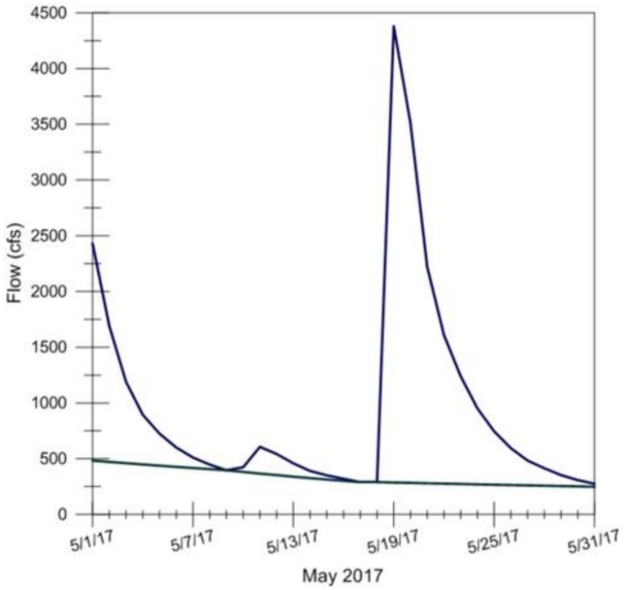


Figure 13 Barren Fork at Eldon, separated hydrograph (streamflow and baseflow) according to HYSEP Local minimum method

PART

The USGS program PART for hydrograph separation is based on antecedent streamflow recession (Barlow et al., 2015) The analysis evaluates if streamflow on N interval of days is greater than or equal to streamflow on the next day within the interval; if this is true baseflow is set equal to streamflow for the day of interest. Linear interpolation is used to assign the baseflow value for the remaining days that did not meet the antecedent recession requirement (Table 9, Figure 13).

Table 9 Example determining baseflow value according to PART method

Day of Interest: May 9, 2017	
Watershed Area: 307 square miles	
Interval: 3 days	
Date	Total Streamflow (cfs)
May 6, 2017	601
May 7, 2017	511
May 8, 2017	448
May 9, 2017	397 (baseflow)

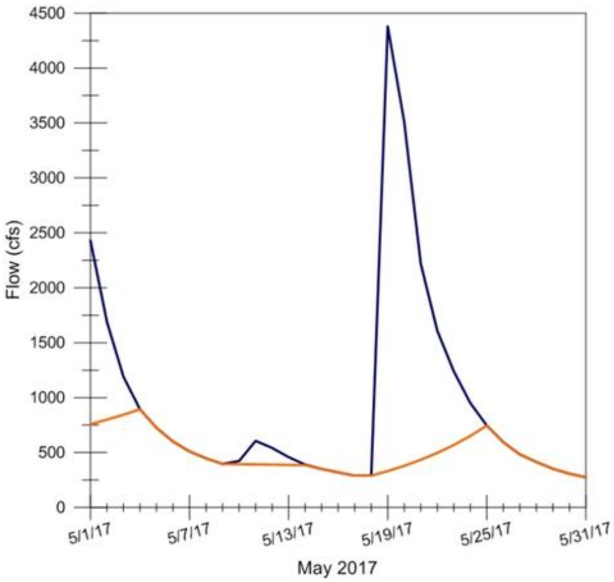


Figure 14 Barren Fork at Eldon, separated hydrograph (streamflow and baseflow) according to PART method

Figure 15 presents the comparison of all three USGS methods for Barren Fork Creek May 2017.

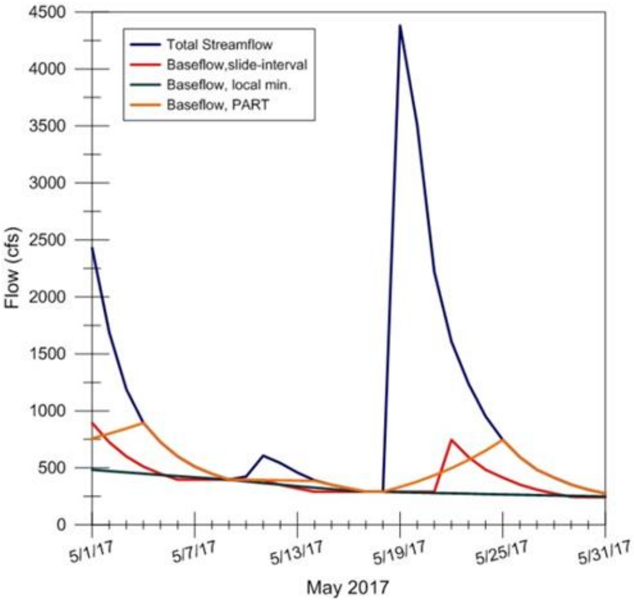


Figure 15 Barren Fork at Eldon, separated hydrograph (streamflow and baseflow) comparison of HYSEP sliding-interval, local minimum, and PART methods

Delta 10%

The final method considered in the analyses is not a hydrograph separation method, but a decision criteria approach to determine if flow conditions on the day of interest are representative of baseflow. This method was based on the approach described by McCarty and Haggard (2016) in a previous analysis in the Illinois River watershed. The Delta 10% method combined two metrics 1) a 48-hour antecedent rainfall benchmark of less than 0.01 inches and 2) a change in daily average flow of less than or equal to ± 10%. When both of these two metrics were met total flow on the day of interest was considered unaffected by direct runoff from a storm event.

Hydrograph Separation Analyses & Results

The critical condition term was defined as the condition when "surface runoff is not the dominant influence of total flow"; thus, a simple analysis was designed to evaluate when direct runoff (i.e., surface runoff) was not the dominant component of total flow. Direct runoff is not dominant when baseflow is dominant; this was interpreted as when baseflow was greater than 50% of the total flow. Increasing percentages of baseflow above 50% were considered to evaluate how a baseflow percentage threshold would impact the availability of eligible water quality sampling days and subsequently the use of associated total phosphorus data in water quality assessments. The objective of the analysis was to identify the population of days that if a sample was collected the result would qualify to assess the total phosphorus criteria.

The hydrograph separation methods described above were applied to the USGS gages in the watershed to evaluate baseflow versus direct runoff. Daily average flows from the gages listed in Table 1 were used in this analysis. The USGS program Groundwater Toolbox provides both a graphical and mapping interface for the analysis of hydrologic data and contains several hydrograph separation methods (Barlow et al., 2015). The eleven-year flow records were uploaded into Groundwater Toolbox and the hydrograph separation programs for HSYEP sliding-interval, HYSEP local minimum and PART were completed. The output for each method included the original daily average flow in cfs, baseflow in cfs calculated according to each method, and percent baseflow. Calculations for the Delta 10% method were performed in Microsoft Excel®.

The analysis was completed for seven locations in Table 1; results for the Illinois River at Tahlequah, Watts, & south Siloam Springs locations are presented below and results for remaining locations are in Appendix 1. As the baseflow percentage threshold increased, meaning baseflow was becoming more and more dominant, the number of eligible sampling days decreased because the occurrence of flow conditions deemed suitable for monitoring becomes more and more restricted (Figure 16). The hydrograph separation method of HYSEP sliding-interval and PART give very similar results until the baseflow percentage (BFP) was greater than 75 (Figure 16). At this point, the PART method becomes more conservative, meaning a greater portion of the total flow was separated into baseflow and thus a greater percentage of eligible sampling days were retained. The Delta 10% method is a binary decision criterion and because both metrics need to be satisfied the number of eligible sampling days was considerably reduced. Additionally, the antecedent rainfall threshold was set low to assure minimal contribution from direct runoff and this worked to make the overall method fairly limiting. The results from these analyses were similar at the Watts and south Siloam Springs locations (Figure 17 and 18).

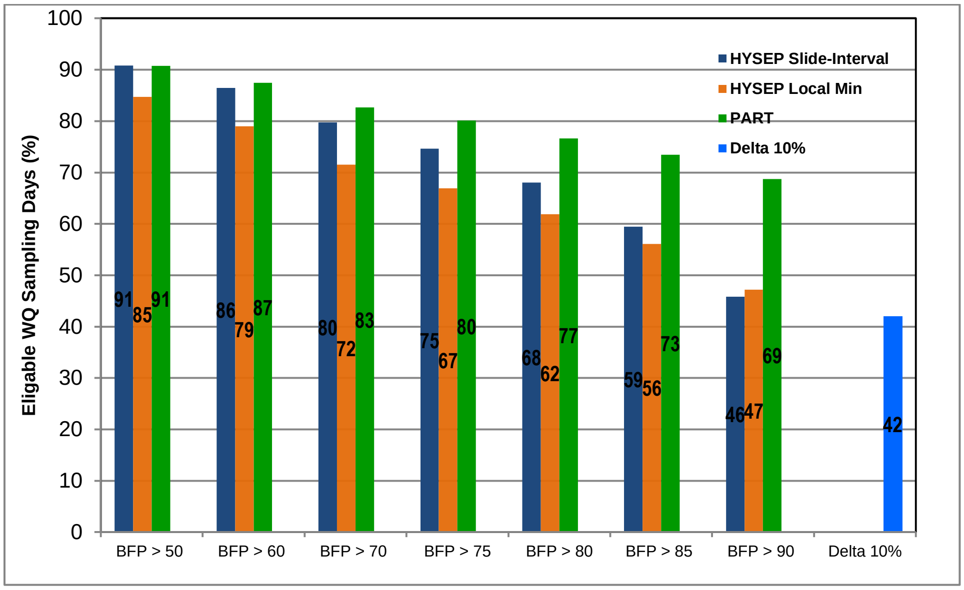


Figure 16 Tahlequah hydrograph separation analysis results; the percent of eligible sampling days decreases with increasing baseflow percent (BFP) threshold.

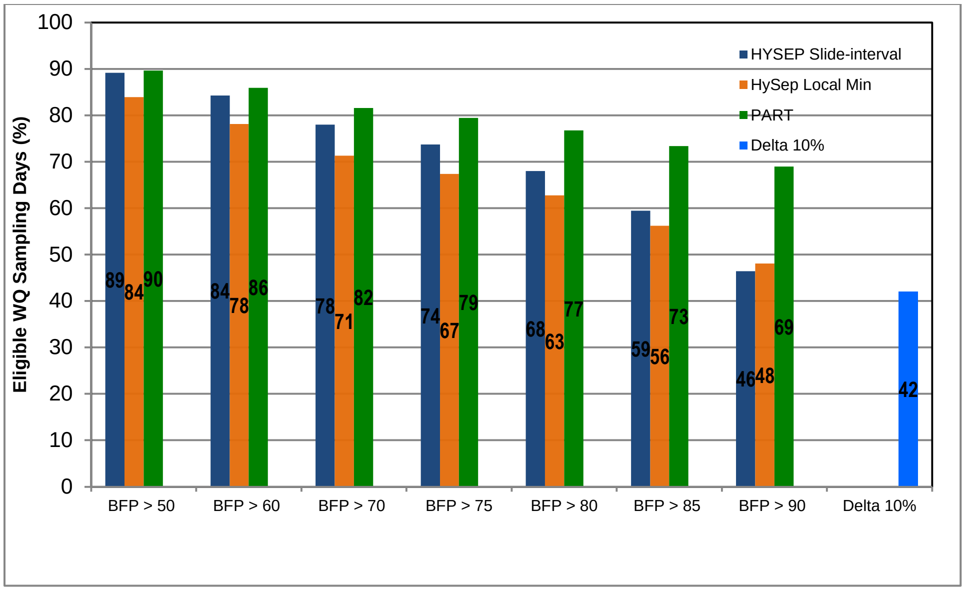


Figure 17 Watts hydrograph separation analysis results; the percent of eligible sampling days decreases with increasing baseflow percent (BFP) threshold.

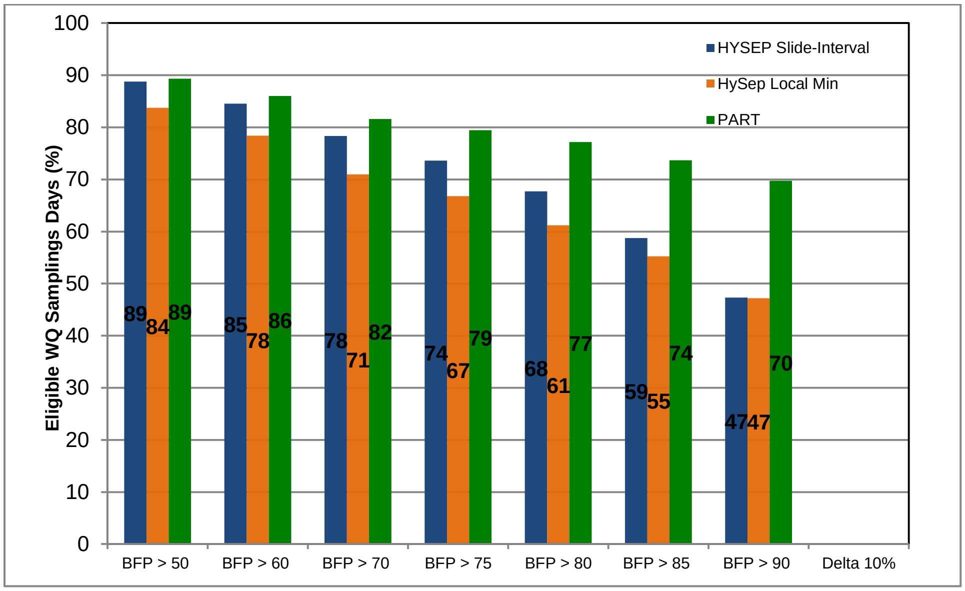


Figure 18 South Siloam Springs hydrograph separation analysis results; the percent of eligible sampling days decreases with increasing baseflow percentage (BFP) threshold.

Upon review of the three hydrograph separation methods and the Delta 10% method, HYSEP sliding-interval was selected for additional analysis. HYSEP sliding-interval was selected because it represented a reasonable compromise between the PART and HYSEP local minimum methods. Additionally, it is an established USGS method and computer programs and

R code are readily available to consistently and efficiently execute the analysis in the future. The Delta 10% method was not selected because comparatively it substantially reduced the number of eligible sampling days and as a binary decision criteria and not a hydrograph separation method it does not allow for further analysis exploring the interaction between different baseflow percent thresholds and the interpretation of ambient water quality conditions. Additionally, the Delta 10% method requires rainfall data and given diverse rainfall patterns and the limited sources for rainfall data this method may prove problematic for consistent long-term implementation. In further analyses on baseflow percentage thresholds and total phosphorus only baseflow from the HYSEP sliding-interval method was evaluated.

Baseflow Threshold Analysis and Results

The critical condition term describes a flow condition considered suitable for collecting data to assess the TP criterion; thus, any given day can be characterized as an eligible or ineligible sampling day based on daily average flow. As presented in Figures 16 - 18, the population of eligible versus ineligible sampling days will vary based on what one determines as a baseflow percentage threshold. The eligible and ineligible sampling days at different baseflow percentage thresholds were displayed on hydrographs for the analysis period of January 2014 through December 2017; a shorter analysis period was used to better view the hydrographs (Figures 19 – 21). This allows for visualization of how the eligible versus ineligible sampling days are represented across flows. The hydrograph is color coded blue and red to reflect the portion of the hydrograph that would be eligible for water quality sampling (blue) and the portion ineligible for water quality sampling (red). The baseflow percentage thresholds of fifty-five percent, seventy-five percent, and ninety percent for Illinois River Tahlequah, Watts, and south Siloam Springs locations are presented below (Figures 19 - 21). As the baseflow percentage threshold increases from 55% to 90%, the portion of the hydrograph eligible for water quality sampling (i.e., the blue portion) becomes increasingly restricted to low flow conditions. The complete series of color coded hydrographs for all locations and various baseflow percentage thresholds are provided in Appendix 1.

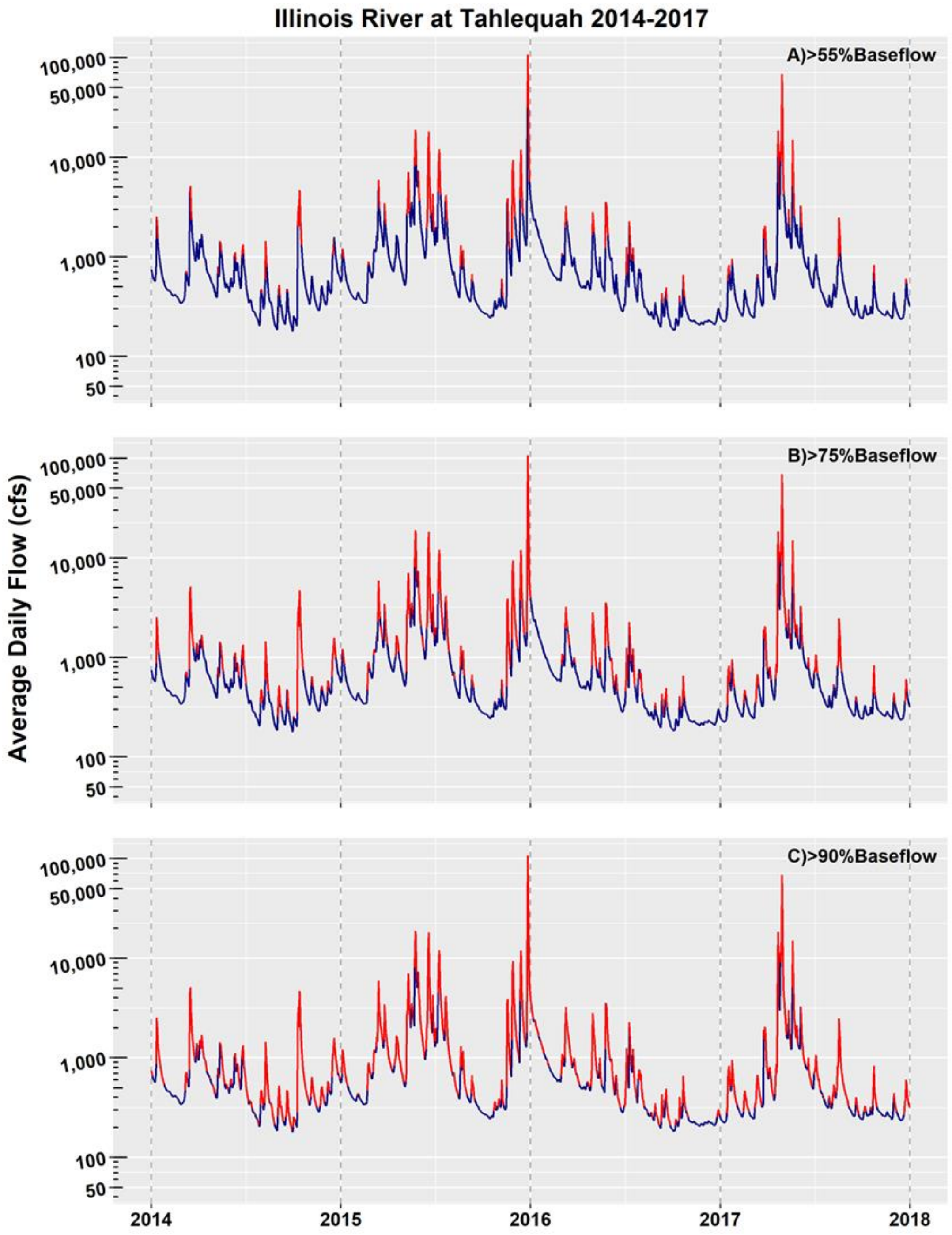


Figure 19 Illinois River at Tahlequah hydrograph (Jan. 2014- Dec. 2017), red indicates portion of hydrograph ineligible for sampling and blue represents portion eligible for sampling, based on HYSEP sliding-interval method and baseflow percentage thresholds of A) 55% or greater, B)75% or greater, and C) 90% or greater.

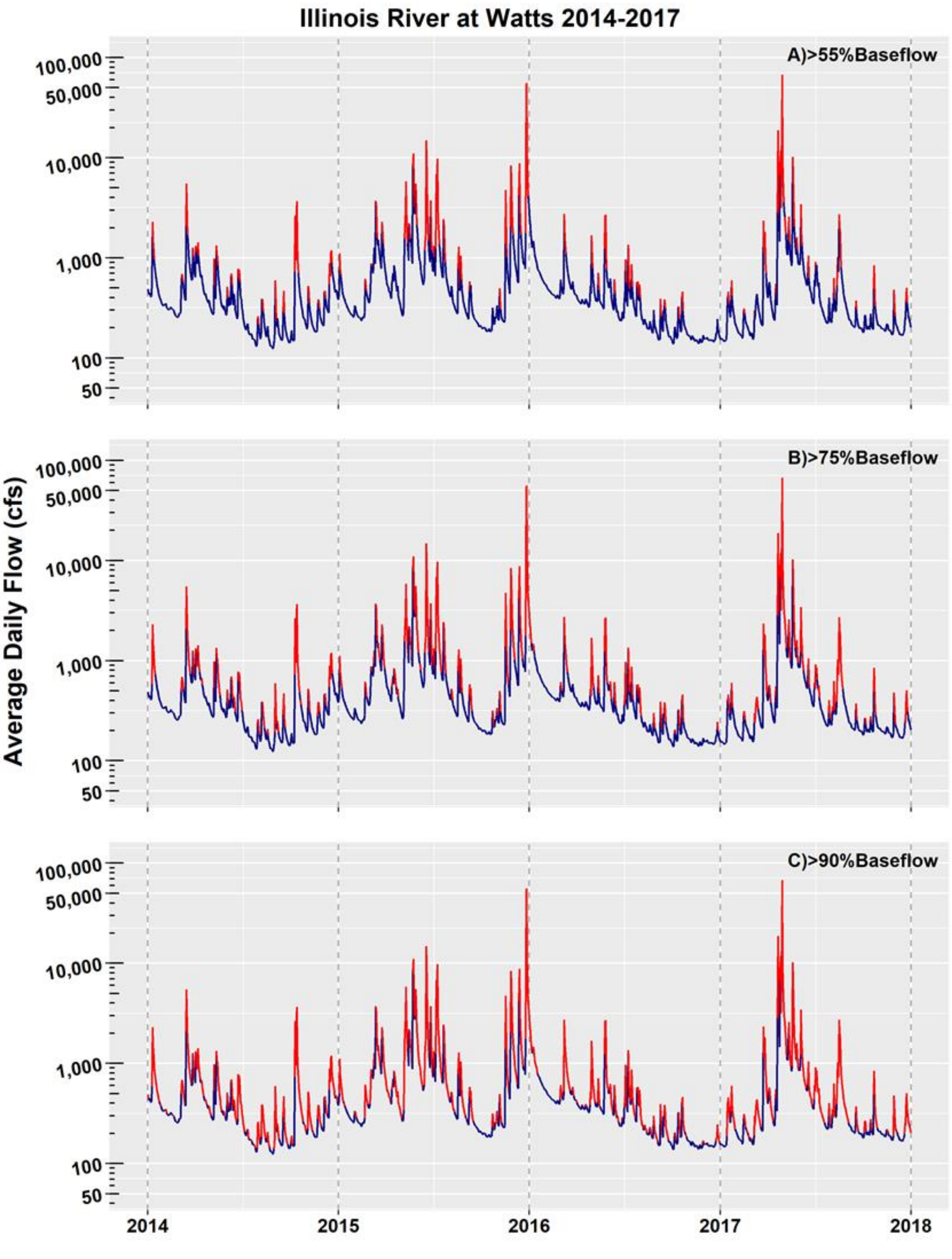


Figure 20 Illinois River at Watts hydrograph (Jan. 2014- Dec. 2017), red indicates portion of hydrograph ineligible for sampling and blue represents portion eligible for sampling, based on HYSEP sliding-interval method and baseflow percentage thresholds of A) 55% or greater, B) 75% or greater, C) 90% or greater.

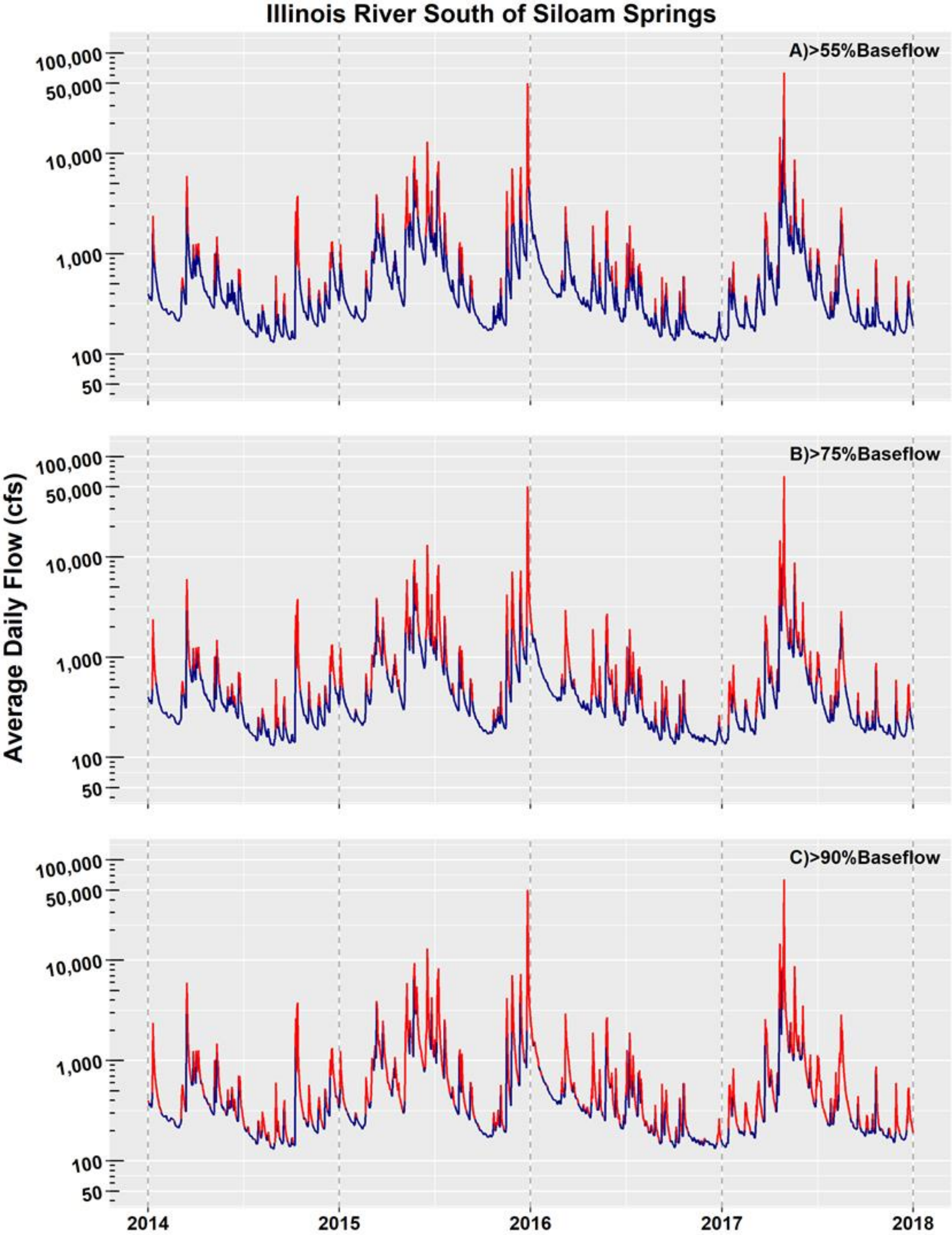


Figure 21 Illinois River at South Siloam Springs hydrograph (Jan. 2014- Dec. 2017), red indicates portion of hydrograph ineligible for sampling and blue represents portion eligible for sampling, based on HYSEP sliding-interval method and baseflow percentage thresholds of A) 55% or greater, B) 75% or greater, C) 90% or greater.

Statistical Analysis

A statistical analysis was conducted to evaluate if the number of eligible sampling days between the different baseflow percentage thresholds was statistically significant. Hypothesis testing of population proportions was conducted. This analysis posed the question: was the number of eligible sampling days at one baseflow percentage threshold significantly different from the number of eligible sampling days at another baseflow percentage threshold? For example, was the number of eligible sampling days at 50% baseflow or greater statistically different than the number of eligible sampling days at 55% baseflow or greater? Table 10 presents the results. All of the results are statistically significant at the alpha of 0.1 or less and by far the majority of results are statistically significant at the alpha of 0.01. There is a strong indication that there is a meaningful difference in the number of eligible sampling days between the various baseflow thresholds.

Table 10 Statistical analysis identifying significance between baseflow percentage thresholds

	Baseflow Percentage Thresholds							
Gaging Station	**50% vs 55%**	**55% vs 60%**	**60% vs 65%**	**65% vs 70%**	**70% vs 75%**	**75% vs 80%**	**80% vs 85%**	**85% vs 90%**
Barren Fork Creek at Eldon	**	***	***	***	***	***	***	***
Illinois River at Tahlequah	***	***	***	***	***	***	***	***
Illinois River at Watts	***	***	***	***	***	***	***	***
Illinois River at South Siloam Springs	***	***	***	***	***	***	***	***
Illinois River at Savoy	***	***	***	***	***	***	***	***
Flint Creek at Kansas	*	**	***	***	***	***	***	***
Osage Creek	*	**	***	***	***	***	***	***
*** significant at <0.01								
** significant at <0.05								
* significant at <0.1								

Baseflow Percentage Threshold & Total Phosphorus

Regular water quality monitoring has been conducted approximately monthly in the Illinois River watershed for about twenty years. Typically, water samples have been collected at the ambient flow on the day of sampling and in accordance with OAC 785:46-15-4(b), which requires a minimum of six storm event sampling occurrences for assessment of the scenic rivers total phosphorus criterion. Implementing a new approach that would only allow data collected when the critical condition was satisfied to be used for the purpose of beneficial use assessment represents a significant transition away from the longstanding Oklahoma monitoring practices and the inclusive use of data for beneficial use assessment. The influence of implementing a critical condition flow on the interpretation of instream water quality was evaluated in the context of both total phosphorus concentration and load.

The monthly total phosphorus data from January 2008 to December 2018 was used to calculate 6-month rolling averages and graphed as a time series. Figure 22 presents the 6-month rolling average total phosphorus concentration at the Tahlequah location. The blue line and triangle symbol represent the 6-month average concentration based on sample values collected at any flow condition and specifically includes sampling of the six storm events. Whereas the brown, green, and orange lines represent the 6-month average TP concentration based on samples that were collected when baseflow comprised greater than 55%, 75%, and 90% of the total flow, respectively. It is clear that restricting the sample results included in the 6-month average based on a baseflow percentage threshold dramatically influences the evaluation of ambient in-stream total phosphorus conditions. The same time series of total phosphorus concentrations at the Illinois River at Watts and south Siloam Springs locations show a similar effect (Figures 23 and 24). As expected, when the baseflow percentage threshold was increased a greater number of sample results continue to be restricted from the 6-month average calculation and the outcome was a calculated lower in-stream TP concentration.

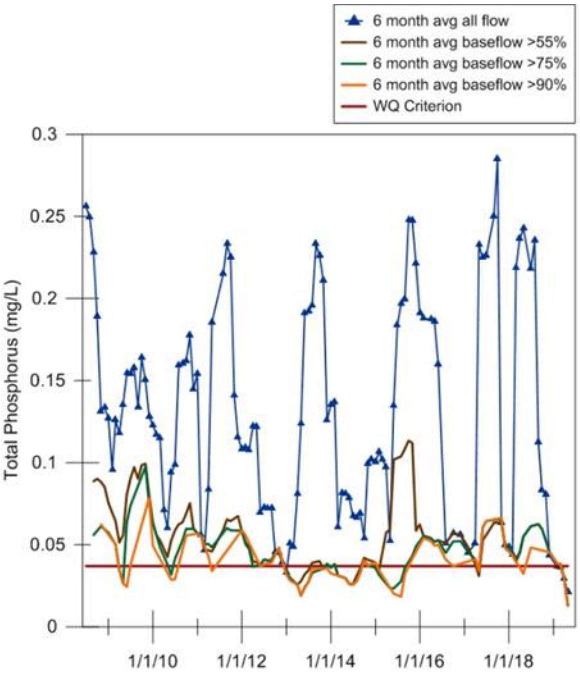


Figure 22 Tahlequah 6-month rolling average TP concentration. The blue triangle symbol represents the 6-month average TP concentration based on sample values collected at any flow condition. Brown, green, and orange lines are the 6-month average TP concentration at > 55%, 75%, and 90% baseflow percentage thresholds. The horizontal red line is the water quality criterion value of 0.037 mg/L.

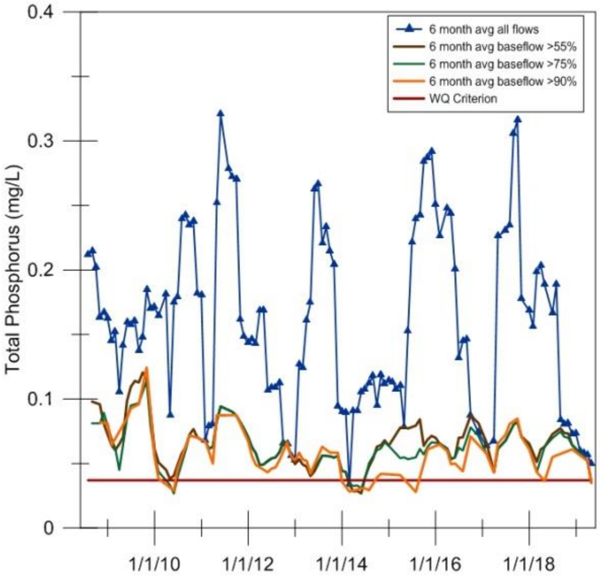


Figure 23 Watts 6-month average TP concentration at all flows and baseflow percentage thresholds 55, 75, and 90 percent. The horizontal red line is the water quality criterion value of 0.037 mg/L.

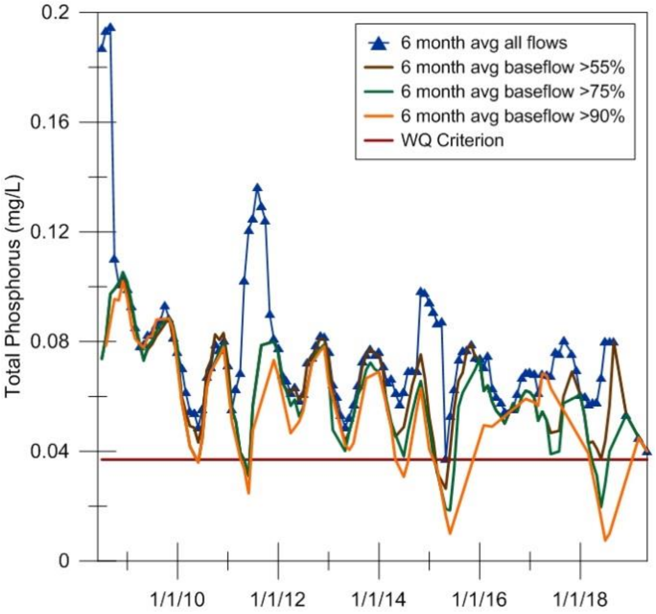


Figure 24 South Siloam Springs 6-month average TP concentration at all flows and baseflow percentage thresholds 55, 75, and 90 percent. The horizontal red line is the water quality criterion value of 0.037 mg/L.

The effect was not so dramatic at the Illinois River at south Siloam Springs location because the ADEQ monitoring program does not specifically include a requirement for storm event sampling (Figure 24). This likely explains why less of a deviation was observed when sample results for the 6-month average TP calculation were restricted to only those collected when baseflow was greater than 55%, 75%, and 90% of total flow.

In-stream phosphorus concentrations are often flow dependent; during wet weather events, direct runoff carries both dissolved and particulate forms of phosphorus directly into the stream. Phosphorus data from the Illinois River watershed typically show greater TP concentrations occurring with increased flow. Therefore, a large portion of the overall phosphorus load is delivered to the stream during wet weather events. Also, the increased flow during wet weather events can activate shallow groundwater flow pathways, which contribute both flow and phosphorus into the stream (Fox et al., 2016, Macrea et al., 2011, Chinnasamy et al., 2014). In contrast, during dry weather periods in-stream phosphorus concentrations are generally lower and contributions are primarily from continuous discharge sources such as wastewater discharges. Understanding the effects of flow on total phosphorus concentration and delivery across a range of flow conditions is important to restoring beneficial uses in the Illinois River, Flint Creek, and Barren Fork creek and protecting downstream waters.

The delivery of phosphorus load across flow conditions can be represented by a load duration curve. Figures 24 - 27 are load duration curves that display how TP load was delivered across the range of flows. A load duration curve is created by multiplying the streamflow by the water quality criterion and required conversion factors. The instantaneous load values in these figures reveal that failure to attain the criterion occurs across all flow conditions; although with a greater magnitude at lower percentile flows.

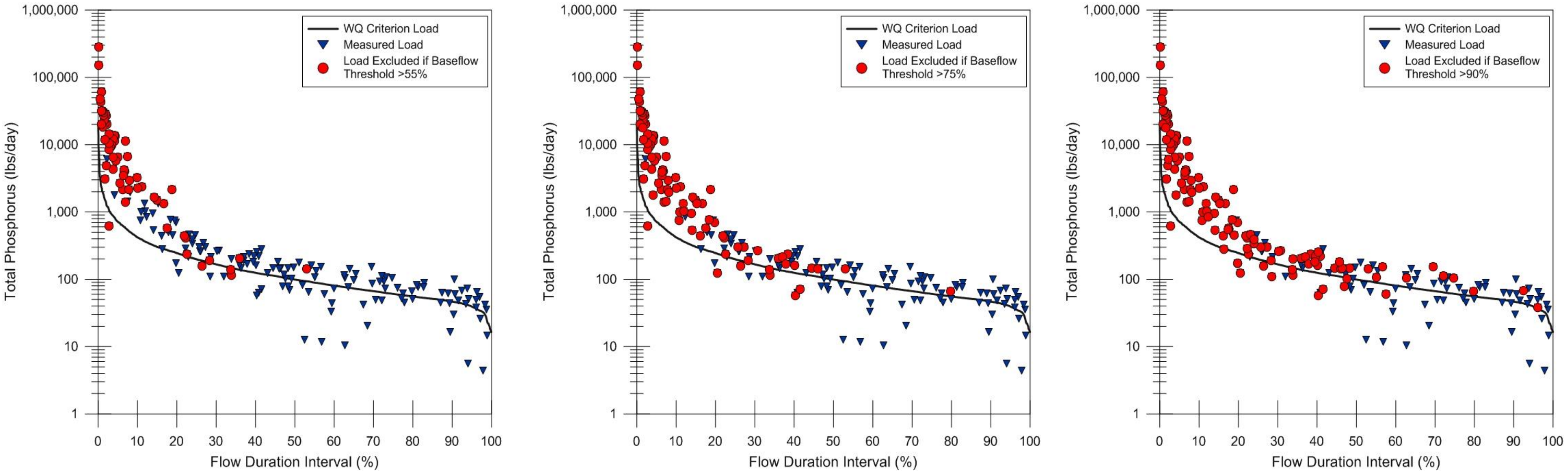


Figure 25 Illinois River at Tahlequah, load duration curve. The solid black line represents the TP load attaining the water quality criterion across flow intervals, and the colored symbols are the instantaneous TP load, based on measured water quality data. Loads that plot above the curve (i.e., black line) indicate an exceedance of the water quality criterion. Red circle symbols indicate the TP load excluded from water quality assessment at the baseflow percentage thresholds of 55, 75, & 90 percent.

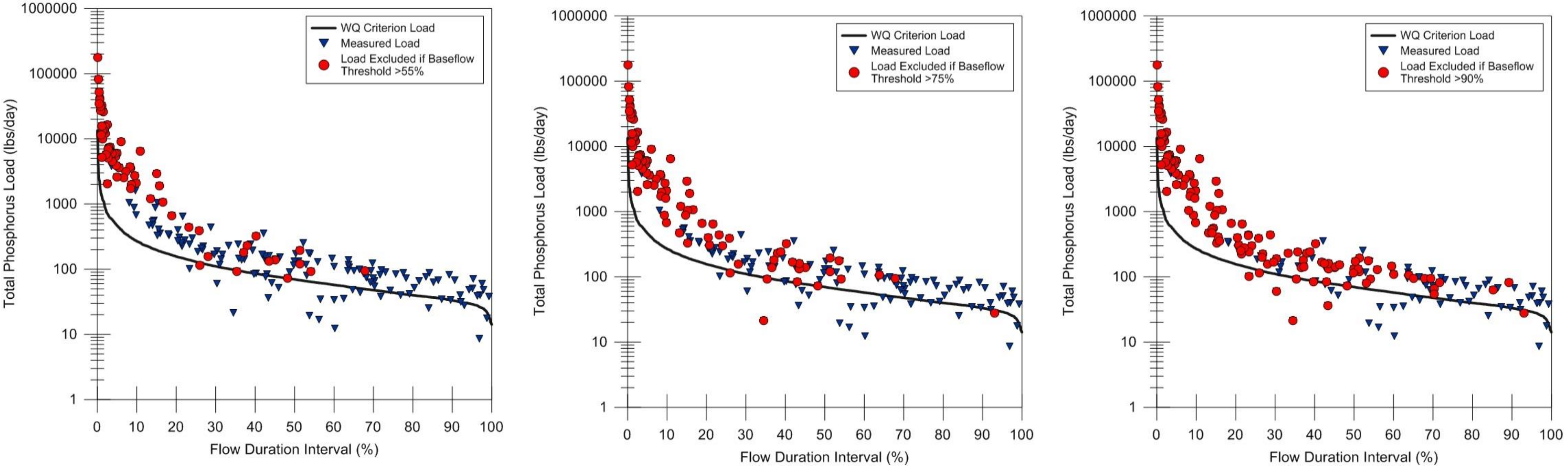


Figure 26 Illinois River at Watts, load duration curve. The solid black line represents the TP load attaining the water quality criterion across flow intervals, and the colored symbols are the instantaneous TP load, based on measured water quality data. Loads that that plot above the curve (i.e., black line) indicate an exceedance of the water quality criterion. Red circle symbols indicate the TP load excluded from water quality assessment at the baseflow percentage thresholds of 55, 75, & 90 percent.

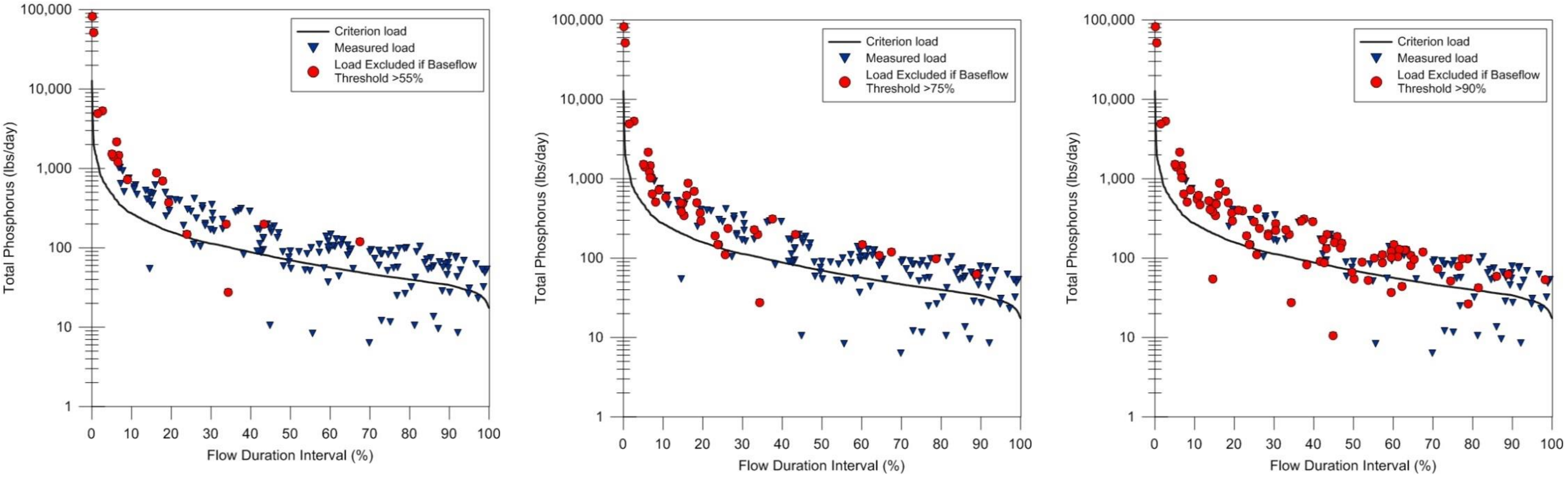


Figure 27 Illinois River at South Siloam Springs, load duration curve. The solid black line represents the TP load attaining the water quality criterion across flow intervals, and the colored symbols are the instantaneous TP load, based on measured water quality data. Loads that plot above the curve (i.e., black line) indicate an exceedance of the water quality criterion. Red circle symbols indicate the TP load excluded from water quality assessment at the baseflow percentage thresholds of 55, 75, & 90 percent.

The red circle symbols represent at which baseflow percentage threshold a load would be excluded from water quality assessment. Working from the prerequisite that baseflow must be 50% or greater of total flow, a substantial amount total phosphorus loading was automatically excluded from consideration at the baseflow percentage threshold of 55% (red circles on figures). As the baseflow percentage threshold increases to 75 and 90 percent it was observed that an increasing portion of the total phosphorus load would be excluded from water quality assessment (Figures 25 - 27). The baseflow percentage threshold dictates which TP loads and consequently which TP sources are represented in a water quality assessment.

This is important because the phosphorus load delivered at different flow duration intervals often comes from different sources. For example, the load delivered to the river at the 80 - 100 flow duration interval represents flows of approximately 100 – 400 cfs, which are non-stormflows and predominately reflect phosphorus contributions from continuous wastewater dischargers. In contrast, the load delivered in the flow duration interval of 1-10 represents flow of 1,100 – 100,000 cfs, which includes stormflows and reflects phosphorus load contributions from nonpoint sources. Thus, as the baseflow percentage threshold increases greater responsibility would be focused on continuous wastewater dischargers to meet the criteria and considerable loads from nonpoint sources would be excluded from the water quality criteria assessment.

Conclusion

Water quality standards are the foundation for water quality protection under the Clean Water Act and set the benchmark for measuring success of various water quality management programs. It is essential that WQS be implementable and functional across programs. Additionally, regulatory equitability is an essential characteristic when developing a WQS because it works to promote collaborative efforts towards pollutant reduction between different sources and management programs.

The application of a baseflow percentage threshold as a means to implement the Joint Committee recommended critical condition language will limit the total phosphorus data that can be used for water quality assessment. It is clear that using flow to set a limitation around data analyses dramatically influences the appearance of ambient TP concentrations in the river and the consideration of loads delivered from the watershed. Yet, the Aesthetic beneficial use in the Illinois River watershed applies at all times and water quality assessment must provide an accurate evaluation of beneficial use condition. In identifying a baseflow percentage threshold

OWRB staff has the objective of maintaining a WQS that is evenhanded and functional for diverse programs and to minimize artificially influencing the view of ambient TP water chemistry. Therefore, based on the considerations of 1) need to evaluate beneficial use condition, 2) influence that flow restrictions have on evaluation of ambient TP concentrations and loads, 3) the need for evenhandedness across water quality programs, 4) loading restrictions under the Antidegradation Policy, 5) foundational science for original criterion magnitude was based on flow weighted TP values, and 6) longstanding monitoring practices, OWRB staff finds that a 55% baseflow threshold would reasonably address the critical condition recommendation from Joint Committee. The 55% baseflow threshold excludes sample results when stormflow is overtly dominating the river and yet aligns with the considerations above.

Scour Analysis Approach

The basic concept for this analysis explored two questions: 1) are velocities capable of producing algal scour occurring near the streambed and 2) can those velocities be related to daily average flow? The critical condition definition proposes the identification of a flow status when surface runoff is not the dominant influence of stream ecosystem processes. Therefore, a technical analysis was needed that would bring together these two endpoints. A scouring flow analysis does this because it was designed to address questions at the intersection of hydraulics and ecosystem process.

The stream ecosystem process of interest in this case was the accrual of benthic algal biomass. Benthic algal biomass was relevant because its growth to nuisance levels is an expression of how the stream experiences stress from total phosphorus pollution. It is also the observable response that directly impairs the Aesthetic beneficial use and the benthic algae *Cladophora glomerata* was the primary response indicator relied upon in the Joint Study (2016) to evaluate an ecologically relevant total phosphorus concentration. The hydraulic endpoint of interest was water velocity. As velocity increases shear stress also increases and when shear stress reaches a critical point the benthic algae will be physically removed (scoured) from the streambed. It is at this point of active scour that the threshold of dominant influence on the ecosystem process of benthic algae biomass occurs.

Analysis

The analysis discussed below evaluates near streambed water velocity and thresholds for algal biomass scour, which allows for a determination of potential scour or non-scour conditions. The scouring velocity values were then related to daily flow and flow increases as a result of surface runoff during rain events. The figure below is a schematic of the analysis approach (Figure 28).

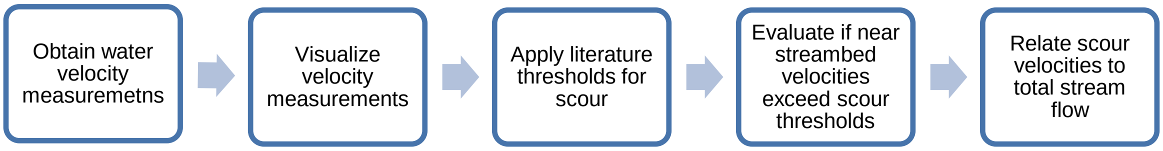


Figure 28 Schematic of conceptual approach for scour analysis

The intention of this analysis was to fully explore the portion of the critical condition definition related to surface runoff as a dominant influence on stream ecosystem processes. Relating scour velocities to stream flow was set as the final aim of the analysis because daily average flow is a readily available metric that would be useful for translating the Joint Study Committee critical condition terminology into an operational definition that could be consistently implemented across programs in both states.

Site Description

This analysis was conducted at two locations Illinois River at Tahlequah (USGS gage 07196500) and Illinois River south of Siloam Springs (USGS gage 07195430). The Illinois River near Tahlequah is located east of Tahlequah near the intersection of Highway 10 and Highway 51 (Figure 29). This site is on the lower end of the upper Illinois River and is the last USGS gauging station on the river before it enters Lake Tenkiller. The substrate at this site was a mix of gravel and bedrock. The ADCP measurements (discussed below) at this site were taken at two locations, the Highway 51 bridge and an old Highway 62 bridge located approximately 1,000 ft. downstream from the Highway 51 bridge. The old Highway 62 bridge was removed in 2017, until that time, measurements were primarily taken from this location when the discharge was below 20,000 cfs. When the discharge was greater than this, access to the road leading to the old Highway 62 bridge was closed and the measurements were taken from the Highway 51 bridge.

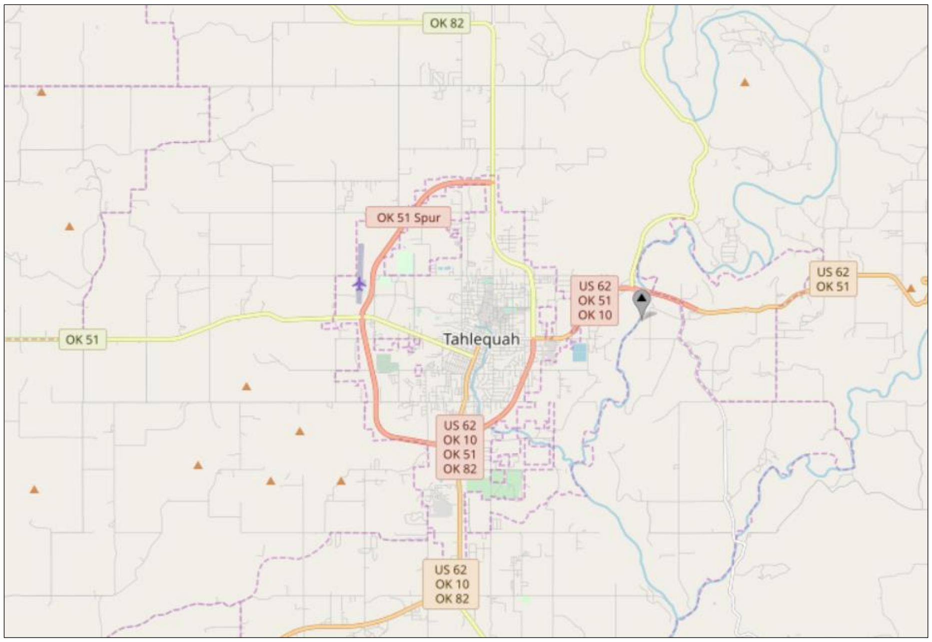


Figure 29 Site location (gray map pin) for Illinois River at Tahlequah.

The Illinois River south of Siloam Springs is located approximately 5 miles south of town on Arkansas Highway 59 (Figure 30). The substrate at this site was a mix of gravel, sand, and bedrock. This is a popular recreation spot with floating/camping outfitters operating on the east and west side of the highway. With the exception of a few measurements taken from a kayak because of USGS staff safety concerns, the majority of ADCP measurements (discussed below) at this location were taken from the Highway 59 bridge.

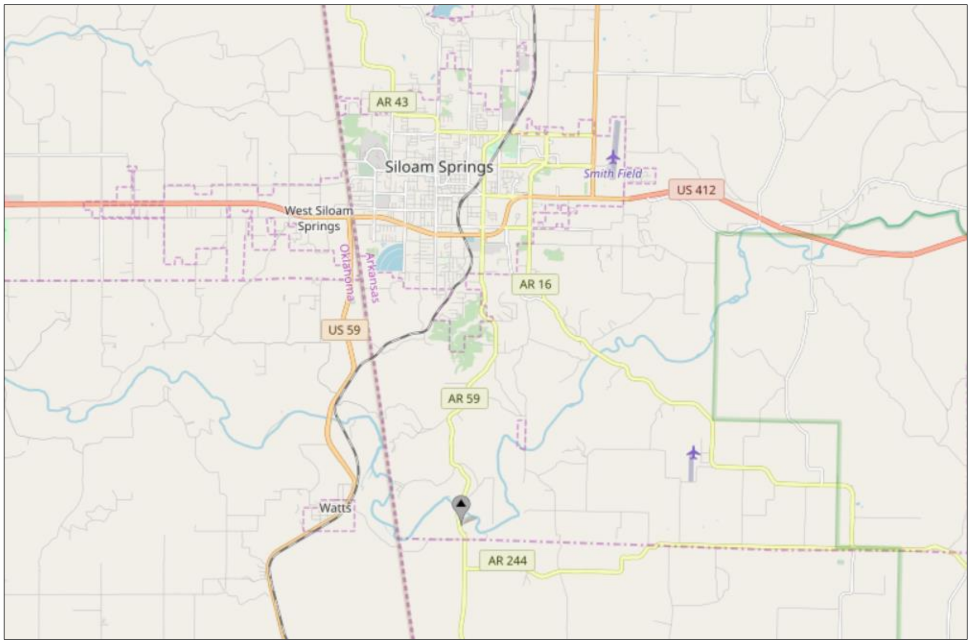


Figure 30 Site location (gray map pin) for Illinois River at south Siloam Springs.

Data Sources & Preparation

The essential data for this analysis were 1) water velocity, 2) velocity thresholds for algal scour, and 3) daily average flow. Water velocity data from 2008-2018 was obtained from the USGS Oklahoma Water Science Center and Lower Mississippi-Gulf Water Science Center. Velocity thresholds for algal scour were based on reported literature values and daily average flow was from the identified USGS gages.

Water Velocity Data

The velocity values were acquired from Acoustic Doppler Current Profiler (ADCP) measurements by USGS staff as part of measuring stream discharge. The ADCP uses sound waves to measure the velocity of particles in the river with the Doppler shift principle. Doppler's principle relates the change in frequency of a source to the relative velocities between the source and the observer. The development of the ADCP has provided hydrographers and hydrologists with a tool that can substantially reduce the time for making discharge measurements and can measure water velocities at a spatial and temporal scale that was previously unattainable. The ADCP has evolved during the last 25 years from an experimental instrument capable of measuring velocity and computing discharge in deep water (greater than 11 feet) to an instrument that is commonly used to measure water velocity and discharge in streams as shallow as 1.0 ft. deep (Christensen and Herrick, 1982; Simpson and Oltmann, 1993; Oberg and Mueller, 2007b). These instruments are used regularly to measure riverine and estuarine water discharge, to collect data for hydrodynamic model calibration and verification, to assess aquatic habitat, and to study sediment transport processes.

An ADCP may be deployed from a boat or tethered from a bridge (Figure 31) and applies the Doppler principle by reflecting an acoustic signal off small particles of sediment and other material present in the water. Typical boat mounted ADCPs have three or four beams pointing between 20 and 30 degrees from the vertical. Three beams are required to obtain a three-dimensional velocity measurement. If a fourth beam is present, an additional quality check can be measured. Measurements taken from a boat use the bottom of the stream or a global positioning system (GPS) as a reference to determine the water velocity relative to the movement of the instrument. ADCPs are called profilers because they provide measurements of velocity throughout the water column (Figure 31). The ADCP divides the water column into depth cells, called bins, and reports a velocity for each depth cell. Traditional point-velocity meters (Rantz, 1982) compute discharge as the product of the cross-sectional area and the mean water velocity perpendicular to the cross-sectional area. The ADCP discharge computation is based on this same approach.

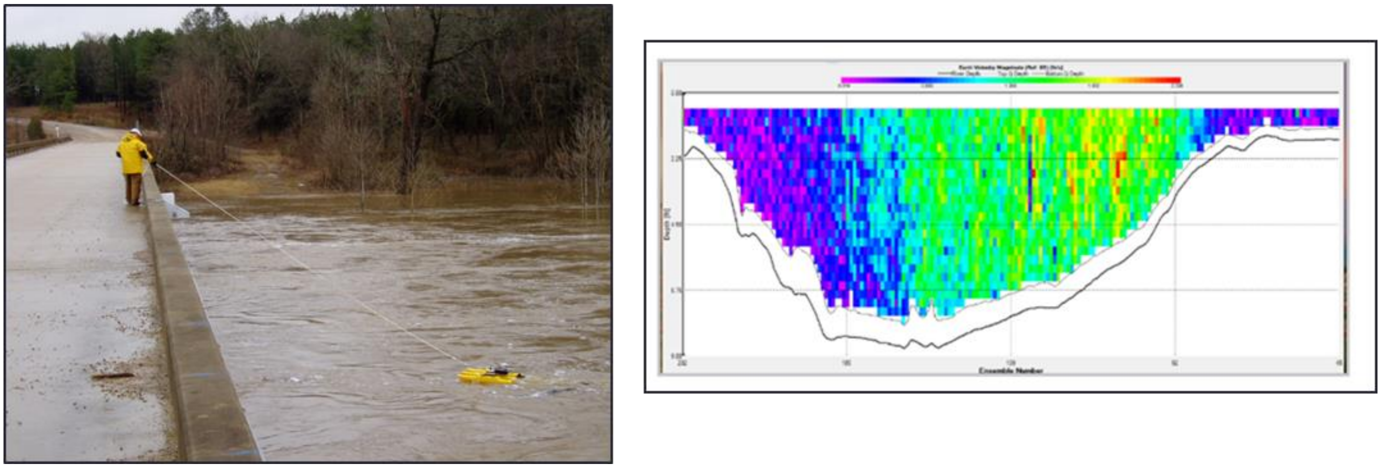


Figure 31 Example of ADCP data collection and transect of velocity profile.

Velocities measured through the water column are not uniform; the maximum velocity occurs at the surface and velocities decrease near the streambed. The velocity values needed for this analysis were those occurring at the bottom of the water column near the streambed because it is at this location that the hydraulic metric (velocity) and the ecosystem process metric (benthic algae) intersect. Because the ADCP measures velocity in bins through the water column creating a profile it was possible to use just the near bed bin velocity value for the analysis. ADCP data from 2008-2018 was requested and obtained from both the Oklahoma and Arkansas USGS offices. The ADCP instruments collect and process data in proprietary software programs. Therefore, the first step to working with the velocity data was to extract the data from the software programs so both velocity profile and individual bin velocity measurements could be obtained. The data was extracted from the software and stored in Microsoft Excel files and further analyses were done in R code.

The ADCP collects profiles in a transect fashion across the stream channel (Figure 31) and under high flow events (those contributing surface runoff to the stream) it is the natural response of the river to rise in stage and spread out into the floodplain. However, it is only the active channel that was of concern in this analysis because it is only in the active channel that under more typical flow conditions nuisance benthic algal biomass grows and impairs beneficial uses. Key preparatory steps in the analysis included exporting data from proprietary software and identifying the active channel and to only evaluate near streambed velocity measurements in this area.

Exporting Data

The ADCP data that was used in this analysis was collected and processed by the USGS. Additional review of the transect velocity profiles and sampling conditions was completed by OWRB staff before exporting the pertinent data into an ASCII file, the format useable in statistical and graphing software. The OWRB staff review of transects had four main goals:

1. Data quality check
2. Visualize the channel and understand the channel dynamics at various flows
3. Information gathering and data export
4. Trimming high flow transects to exclude areas outside the channel

Each of the measurement files were viewed in order to confirm sampling date and any site information notes that were recorded by the field staff and/or processor. The site notes provide valuable information such as exact measurement location, site conditions, confidence in measurements, the type of ADCP used, and any technical issues with instrument that may affect the data.

The discharge summary information was used to check for consistency within a measurement, specifically similarity in channel width and discharge across all transects in a measurement. The velocity profiles were viewed to assess the transects for consistency and completeness, as well as, any "noise" in the transects caused by debris or loss in signal that show up as gaps or blank spots in the profile that need to be flagged or thrown out. If the discharge of an individual transect is greater than +/- 5% from the mean measurement discharge, that transect was not included in the analysis.

Once all ADCP data for a given measurement was confirmed for quality and completeness, the data was exported to be used in later analyses. In the proprietary software, data export was completed via ASCII output creation, where a template that the user creates can be loaded to a particular measurement file and all of the data exported will include the same information in the same format. For this project, the ASCII output files included the following parameters: processed ensemble data (a group of velocity measurements); ensemble number, depth, velocity, discharge, and max water depth. The output file is used to pull data from to run the code and create the graphs for analysis.

Trimming Data

Following the initial review and export of the discharge data, high flow velocity profile transects were trimmed to exclude ensembles from outside the active channel. A combination of field experience and best professional judgement was relied upon to identify the active channel and trim the velocity data accordingly. The following steps describe the process.

1. Review of site images and flow data (multiple dates for each site) along with extensive discussions with experienced field staff on channel morphology and river behavior under various flow conditions
2. Field visit to become familiar with the sites
3. Extensive review of flow transections to gain understanding of how transects change under various flow conditions
4. Additional conversations with experienced field staff
5. The active channel area of the data transect was identified
6. Velocity data was trimmed to only include measurements from the active channel area

Once these steps were completed the velocity thresholds for algal scour, described below, were applied to areas of the channel expected to experience scour.

Literature Velocity Thresholds for Algal Scour

High flow disturbance events lead to the loss of benthic algal biomass due to three key mechanisms 1) direct removal as a result of increased shear stress caused by increased water velocity, 2) abrasion from mobilized sediment that physically scrubs algae from substrate, and 3) mobilized gravel/cobble substrate that scrapes and breaks algae off the substrate (Katz et al., 2018). Benthic algae resistance to scour is influenced by factors including the algal physiological features and the length of time since the last scour event. Benthic algae with a low vertical profile and strong adhesion are resistant to scour. (Stevenson et al., 1996). Cladophora can be successful at resisting shear stress because it has well-attached holdfasts and a tough thallus that becomes more streamlined at increased flow (Dodds and Gudder, 1992). However, there is a tipping point at which increased velocity along with mobilized substrate will remove benthic algae (Francouer and Biggs, 2006 + others).

It is this velocity tipping point that was of interest in this analysis; this is the point at which stream hydraulics dominates the ecosystem process of benthic algal biomass. The paper by Flinders and Hart (2009) conducted a mesocosms study identifying velocities that lead to scouring of benthic algal biomass. In this study algal assemblages (numerically dominated by diatoms and *Cladophora glomerata*) within the mesocosm channels were subjected to velocities ranging from 20 to 240 cm/s. In mesocosm channels subjected to velocities of 150 cm/s or greater there was a significant reduction in periphyton biomass measured as chlorophyll *a*. The three greatest treatment velocities of 150 cm/s, 180 cm/s, and 240 cm/s all had significant results and were employed as near-bed velocity thresholds for scour in this analysis.

Results

As an example of the initial visualization of the ADCP velocity measurements within the river channel, Figures 33 - 34 below present velocity measurements at different flows for the Illinois River at Tahlequah. The top image is the channel transect with velocity values represented by the continuous color scale. In the bottom image the lowest literature endpoint for algal scour 4.92 ft/s (150 cm/s) was set as a threshold value to simply display the potential occurrence or nonoccurrence of scour. The instantaneous flow at the site in Figure 32 was 1,189 cfs, which is near the long term mean flow of 929 cfs.

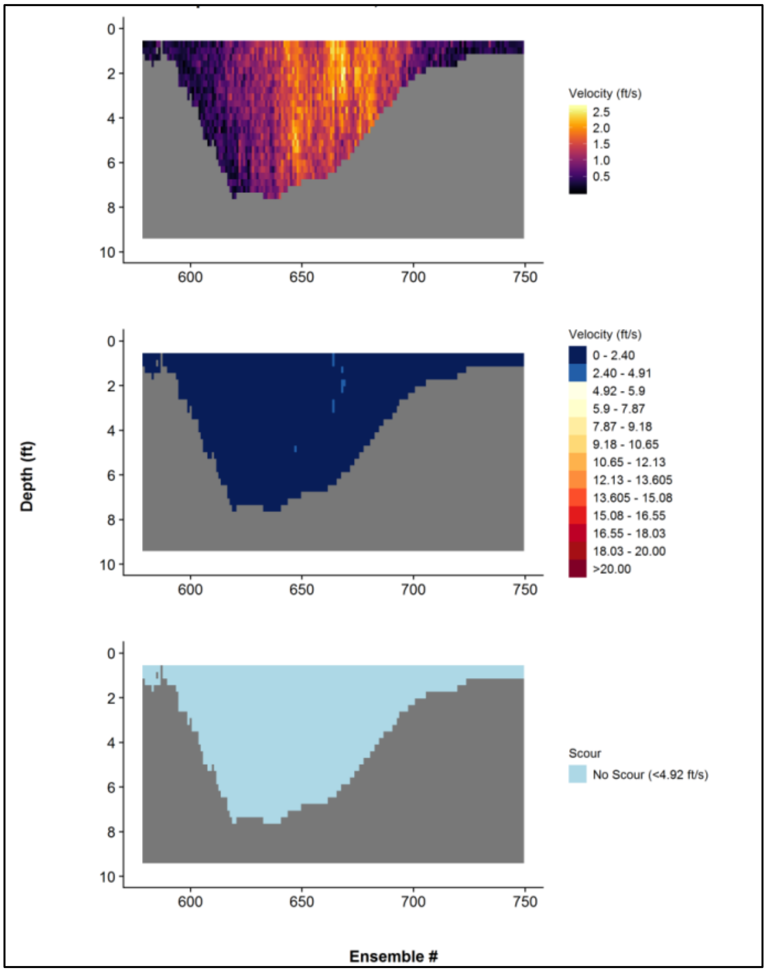


Figure 32 Illinois River at Tahlequah transect profile of ADCP velocity measurements. Top two images display velocity values and bottom image displays presence or absence of scour velocities (flow 1,189cfs).

In comparison as the flow increased almost by an order of magnitude to 9,551 cfs an increase in velocity was observed and the channel has extended toward the left bank (Figure 33). In the lower image additional values for algal scour have been included: scour equals 4.92 ft/s (150 cm/s), intermediate scour equals 5.9 ft/s (180 cm/s) and maximum scour equals 7.87 ft/s (240 cm/s).

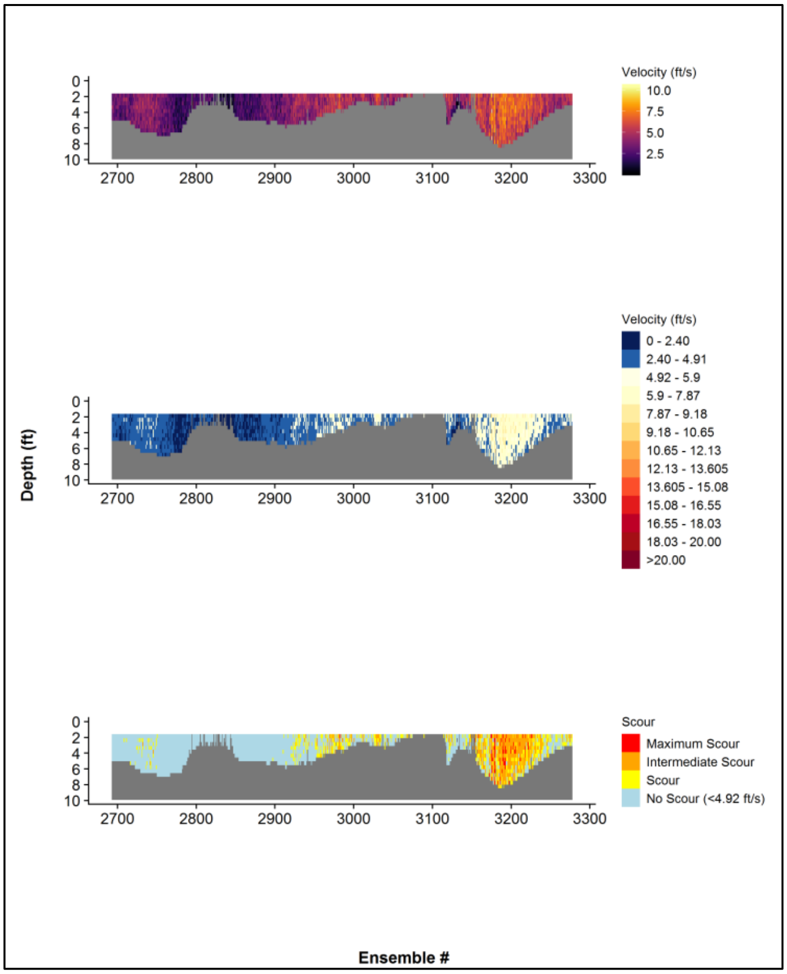


Figure 33 Illinois River at Tahlequah transect profile of ADCP velocity measurements. Top two images display velocity values and bottom image displays presence or absence of scour velocities (flow 9,551 cfs).

Finally, at an extreme high flow event, 85,726 cfs, it was clear that velocities were increased even further and greater areas of the channel are experiencing velocities sufficient to scour benthic algal biomass (Figure 34).

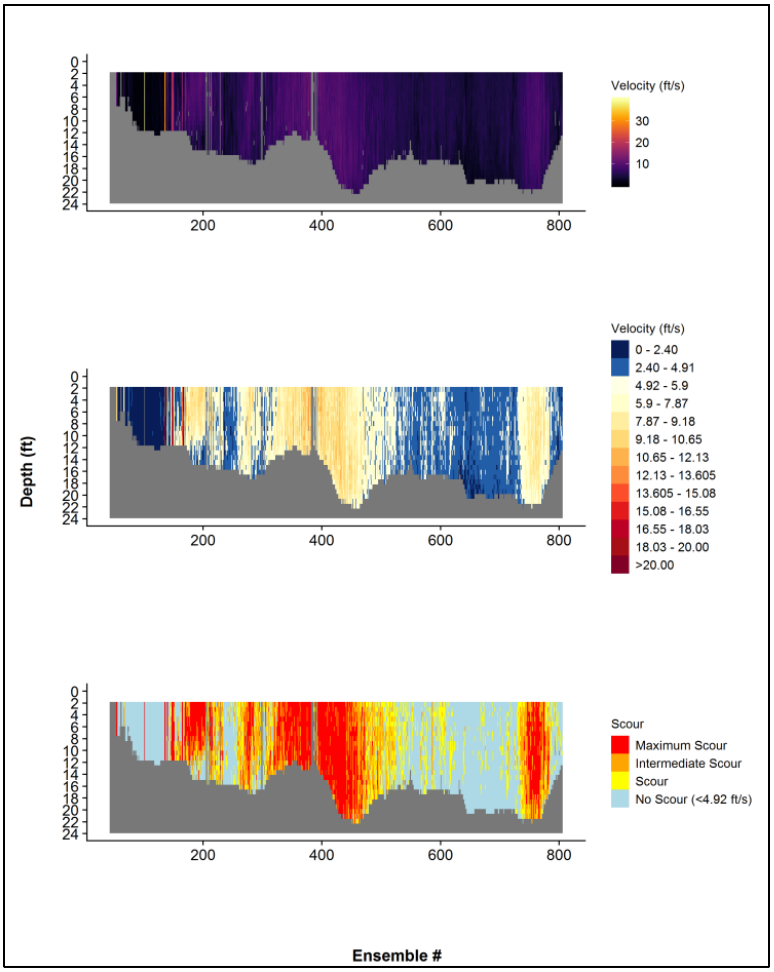


Figure 34 Illinois River at Tahlequah transect profile of ADCP velocity measurements. Top two images display velocity values and bottom image displays presence or absence of scour velocities (flow 85,826 cfs).

The first step in the analysis was to isolate the near the bed velocity measurements within the active channel and evaluate the potential occurrence or nonoccurrence of scour at the streambed. Once again, looking at the Tahlequah site and using similar flow conditions as the previous graphs, it was observed that at approximately 1,100 cfs no scouring velocities were observed near the streambed (Figure 35).

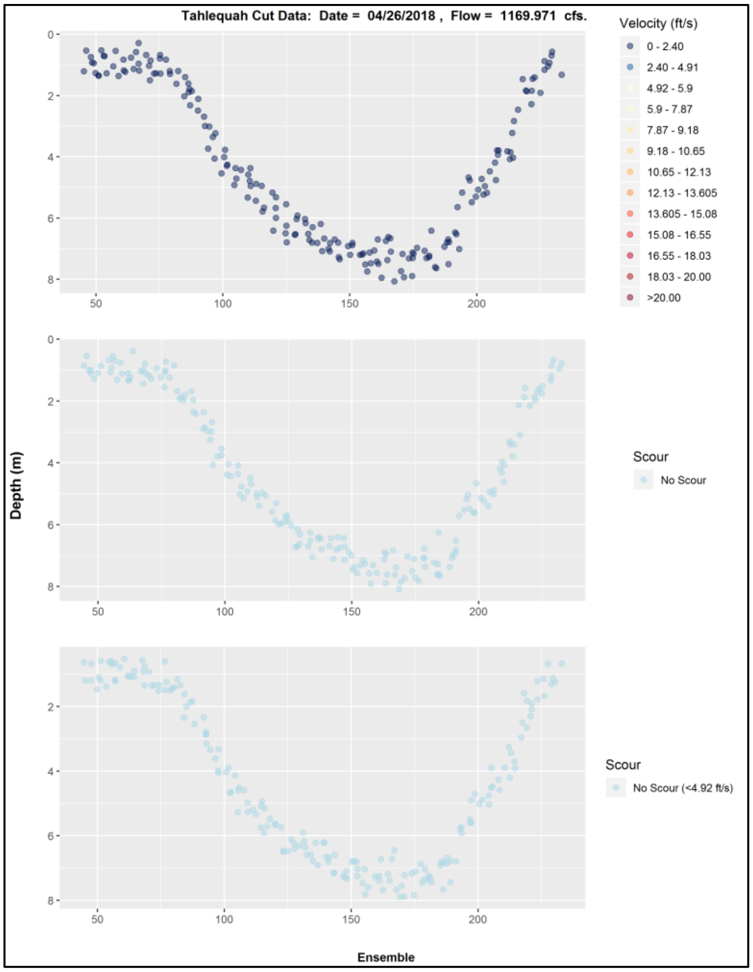


Figure 35 Illinois River at Tahlequah transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 1, 170 cfs).

As the flow increased (~9,500 cfs) near-bed velocities also increased and it was clear that areas of the channel were subject to scouring velocities (Figure 36). As flow and velocity continued to increase (Figures 37 and 38) scouring velocities also continued to be observed; however, it also becomes clear that there was not a consistent pattern or area of the channel subject to scour. Figure 39 presents the scouring velocities when the flow at the site was greater than 100,000 cfs.

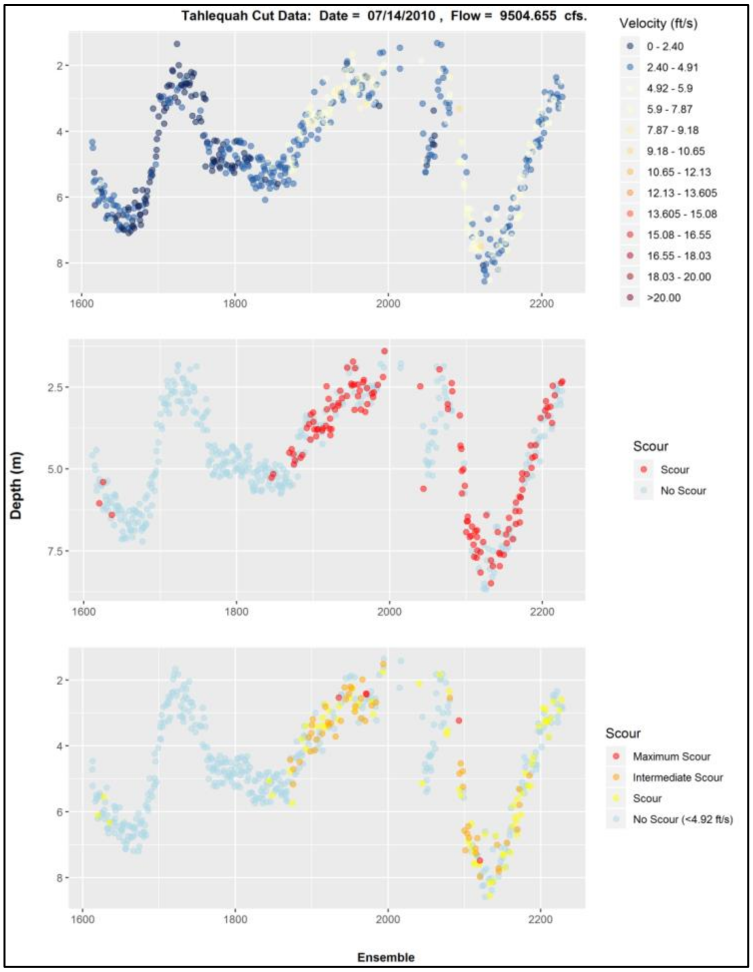


Figure 36 Illinois River at Tahlequah transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 9,504 cfs).

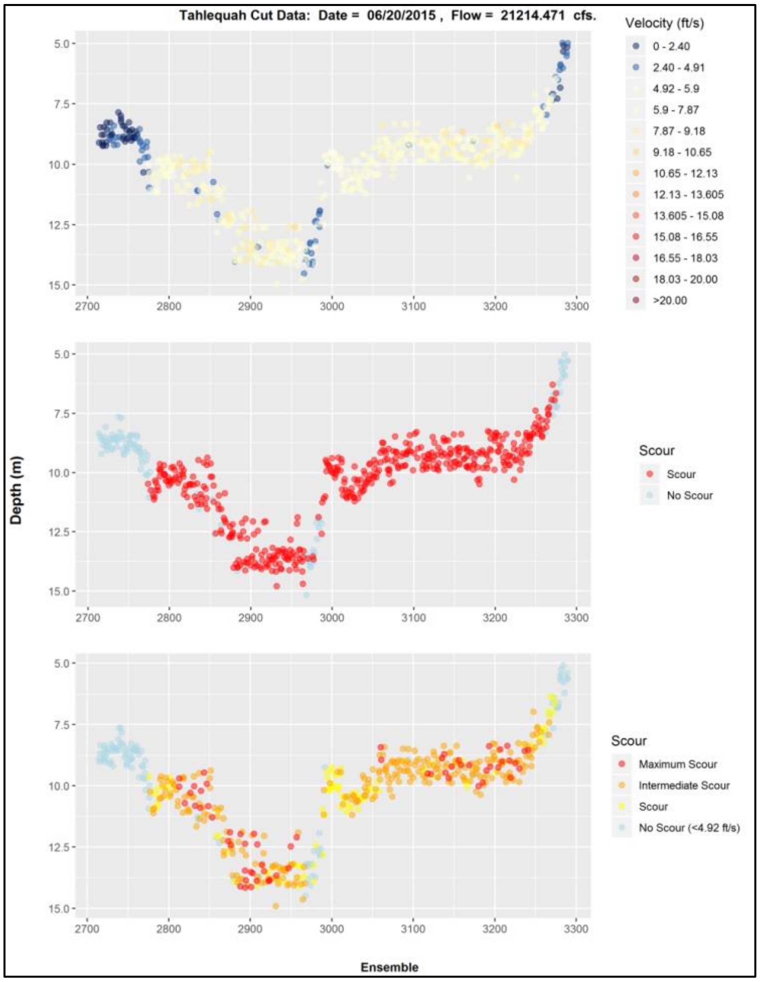


Figure 37 Illinois River at Tahlequah transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 21,214 cfs).

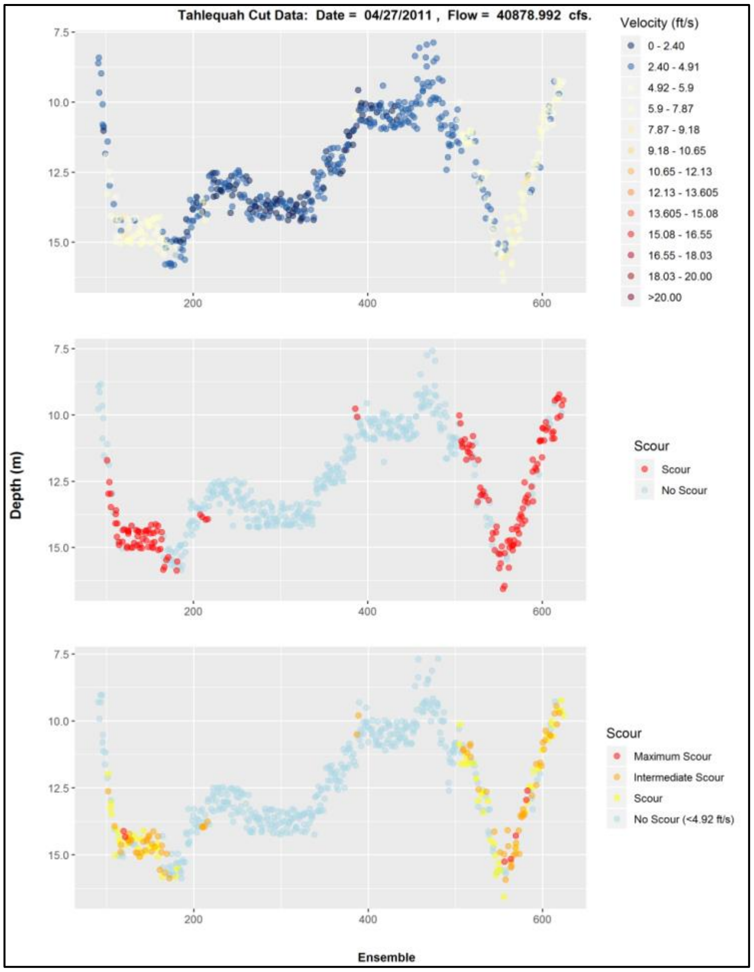


Figure 38 Illinois River at Tahlequah transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 40,878 cfs).

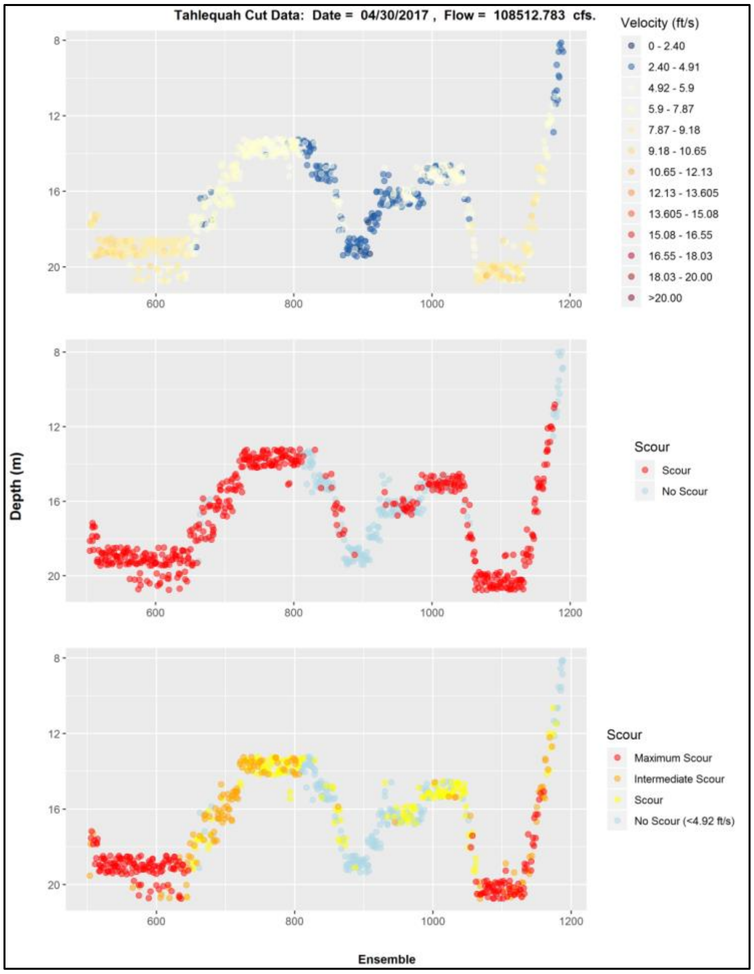


Figure 39 Illinois River at Tahlequah transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 108,512 cfs).

Scouring velocities were not seen at the Illinois River at south Siloam Springs site at a flow of approximately 400 cfs. The long-term average flow at this site is 537 cfs. Figures 41 and 42 are both at a flow of approximately 850 cfs, but from different years and while the frequency of scouring velocities was generally low Figure 42 has none and Figure 41 does have some. This indicates that very similar flow conditions may or may not produce scouring velocities. As flow increases much more substantially the extent of scouring velocities also increased for this site (Figure 43 and 44).

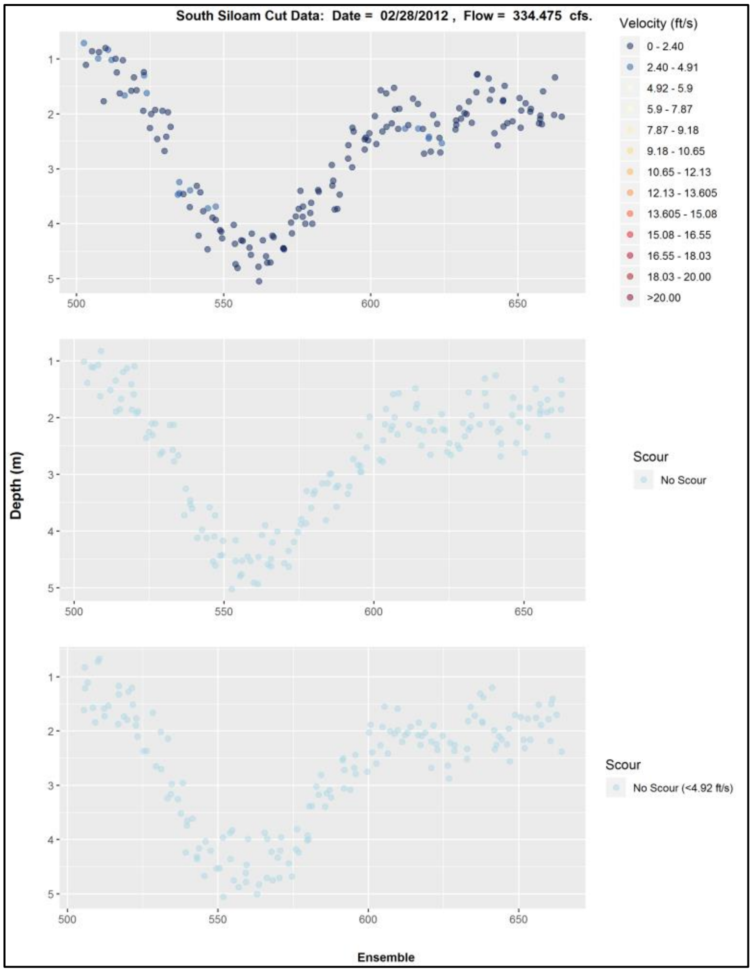


Figure 40 Illinois River at south of Siloam Springs transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 334 cfs).

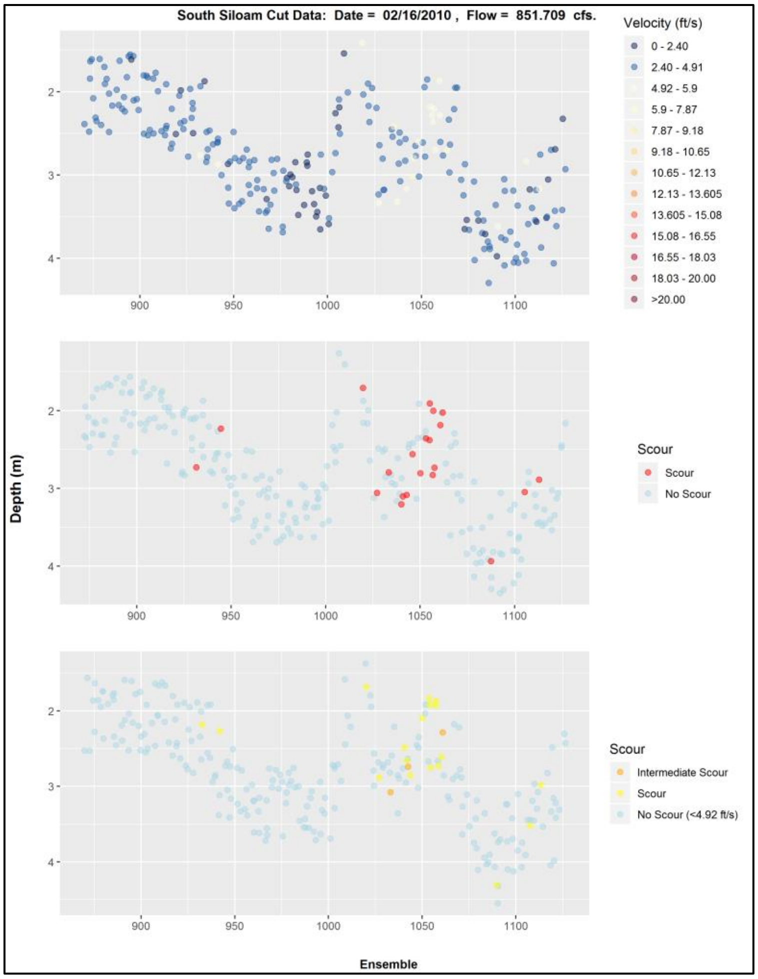


Figure 41 Illinois River at south of Siloam Springs transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 851 cfs).

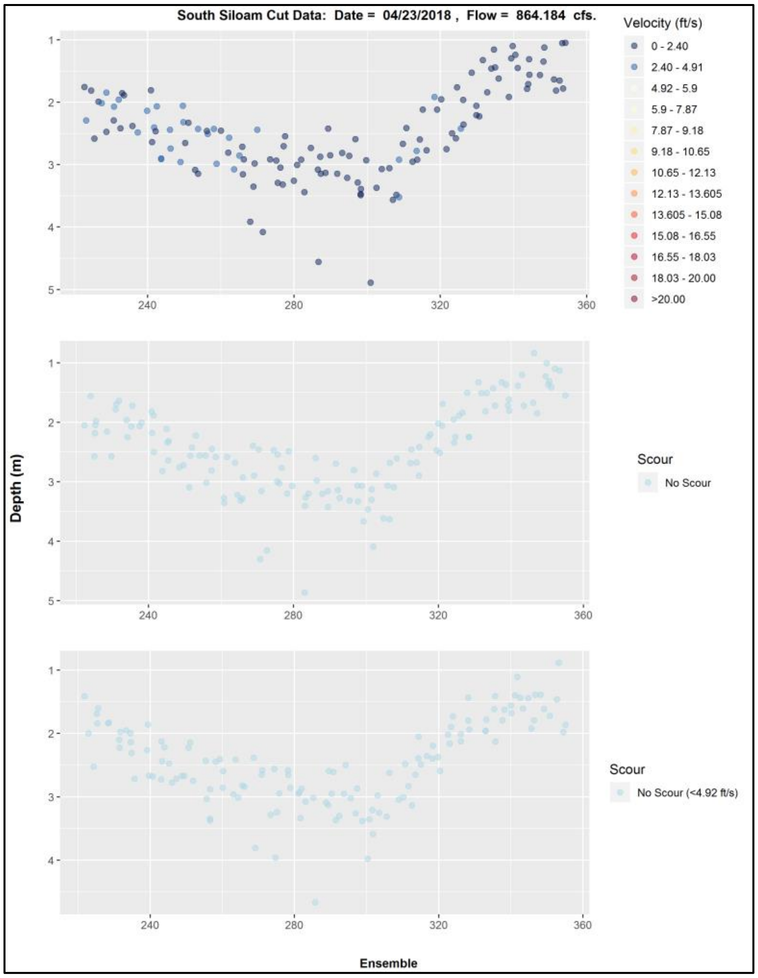


Figure 42 Illinois River at south of Siloam Springs transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 864 cfs).

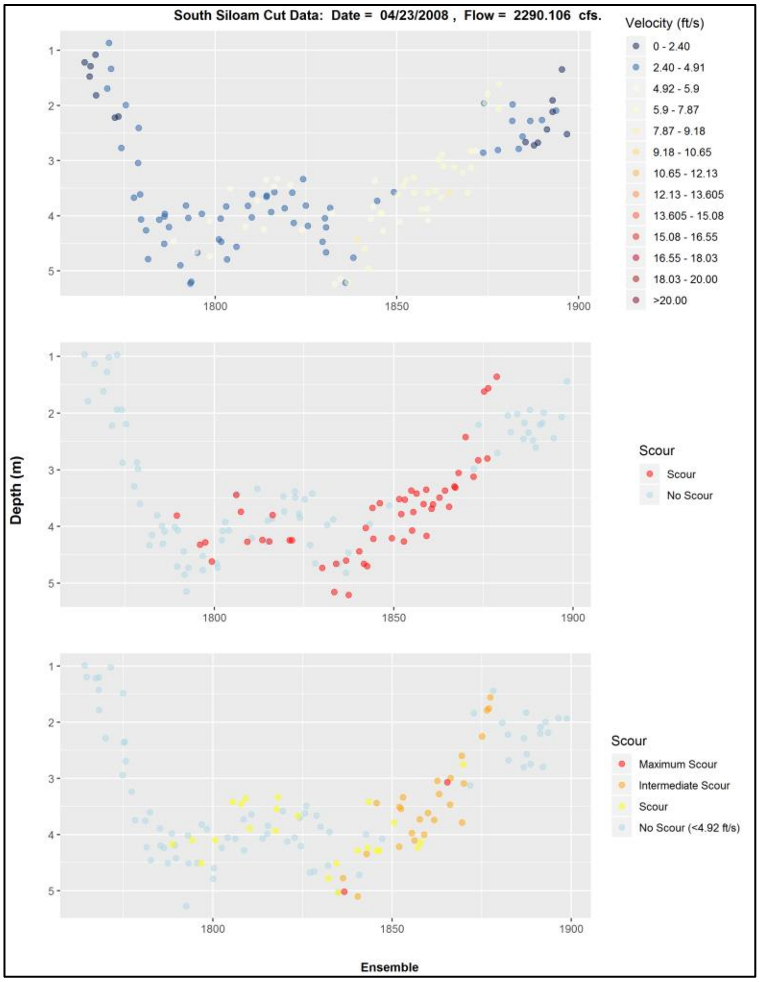


Figure 43 Illinois River at south of Siloam Springs transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 2,290 cfs).

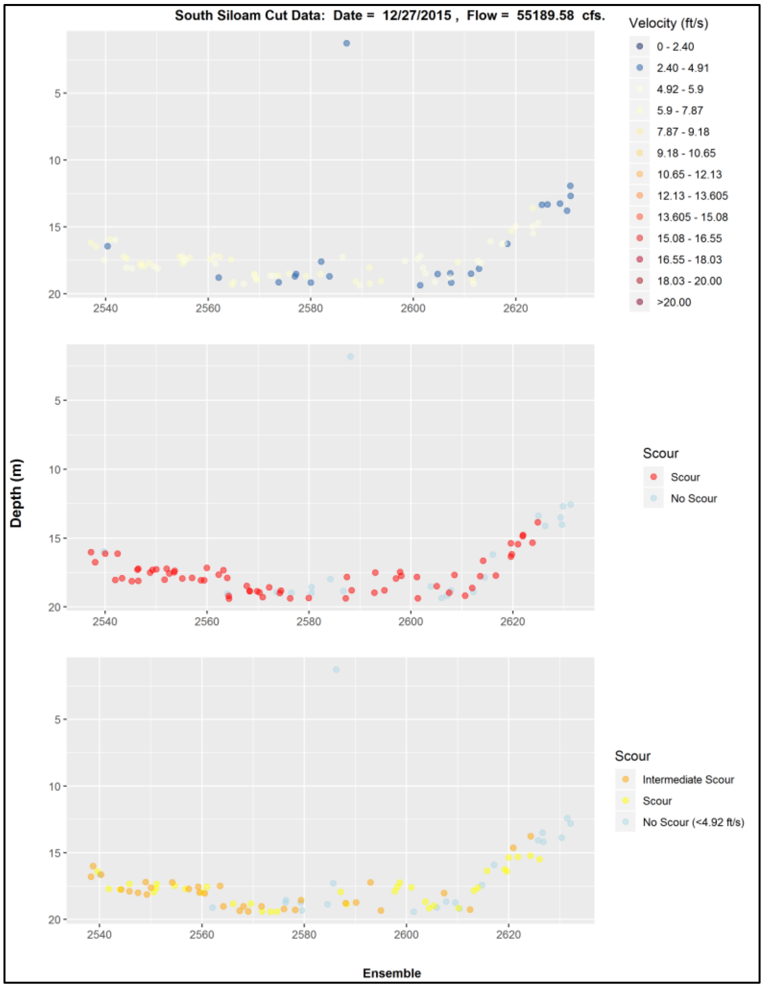


Figure 44 Illinois River at south of Siloam Springs transect of near streambed velocity measurements. Top image display velocity values middle and bottom image display presence or absence of scour velocities according to literature thresholds (flow 55,189 cfs).

The final step in the analysis was to relate near streambed scour velocities to stream discharge. The ADCP measures 100s to 1,000s of individual velocities and the flow gages only report one daily average discharge. Therefore, the near streambed scouring velocities were lumped into daily percent scouring velocities. The figures below relate percent near streambed scouring velocities and daily average discharge for each site.

As previously described, over the years USGS has measured flow for the gage at Tahlequah at two different locations. The stream morphology at these locations varies greatly and affected the flow/velocity relationships and the scouring potential. When the river flows under the Highway 51 bridge overpass, it is wide and flat, the channel is braided and there is a large split on the south or downstream side of the bridge with a large vegetated gravel bar in the middle of the two mainstems (Figure 45). At the old Highway 62 bridge, the channel is no longer braided; it is deep and incised and can contain large volumes of water before it tops the typical stream channel.

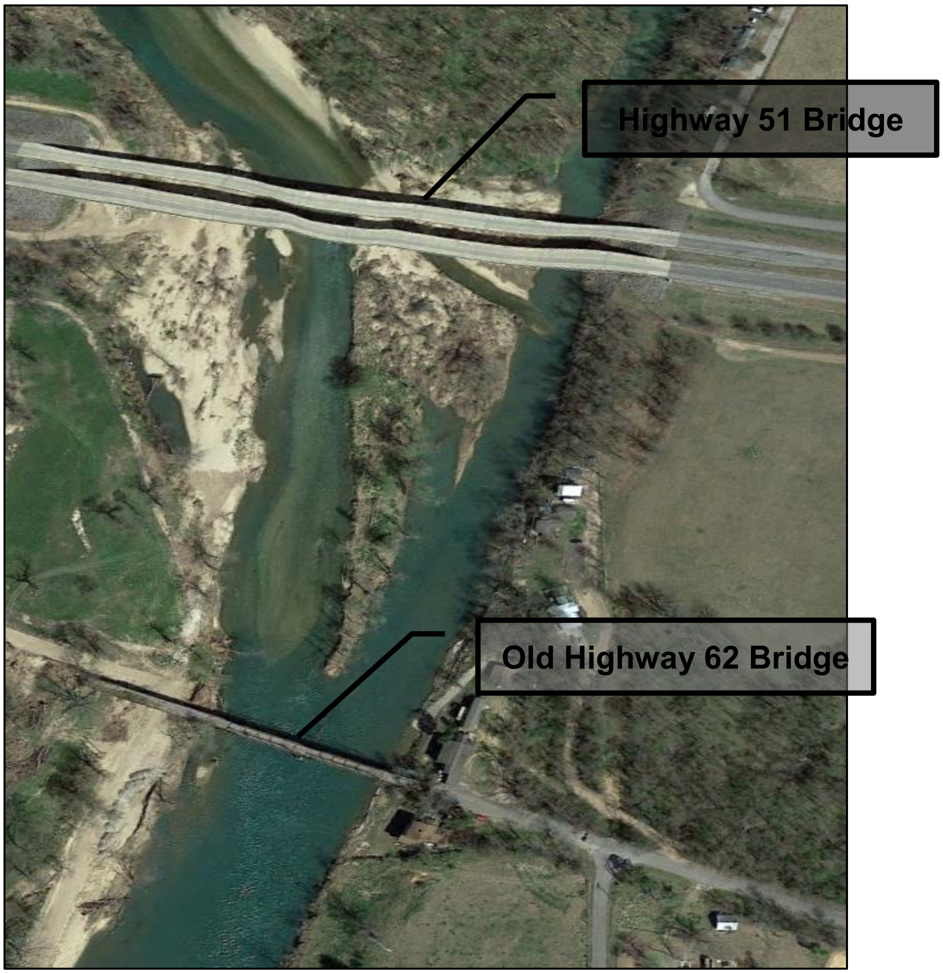


Figure 45 Illinois River at Tahlequah, locations for USGS flow measurements.

When the data from the two different locations are split strong relationships between percent souring velocities and flow were observed and the relationship was in the direction expected (Figure 46). As flow increases the percentage of scouring velocities also increased. However, all of these data were associated with the same gage and when the data were combined the strength of the relationship was considerably diminished. At the south Siloam Springs site considerable variability in percent scouring velocities was observed for very similar flows (~ < 1,000 cfs) (Figure 47).

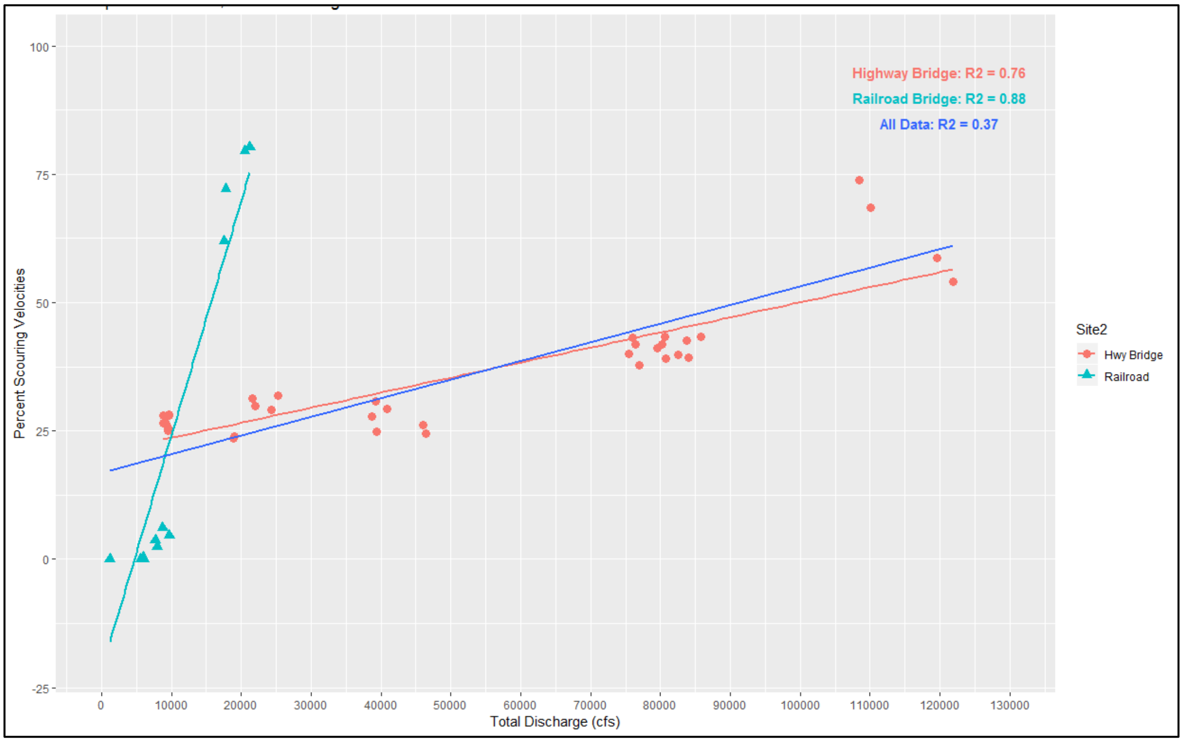


Figure 46 Illinois River at Tahlequah, relationship between percentage of scouring velocities and total daily average discharge. Green line and symbols, old Highway 62 bridge (not railroad bridge), red line and symbols, Highway 52 bridge, blue line data combined.

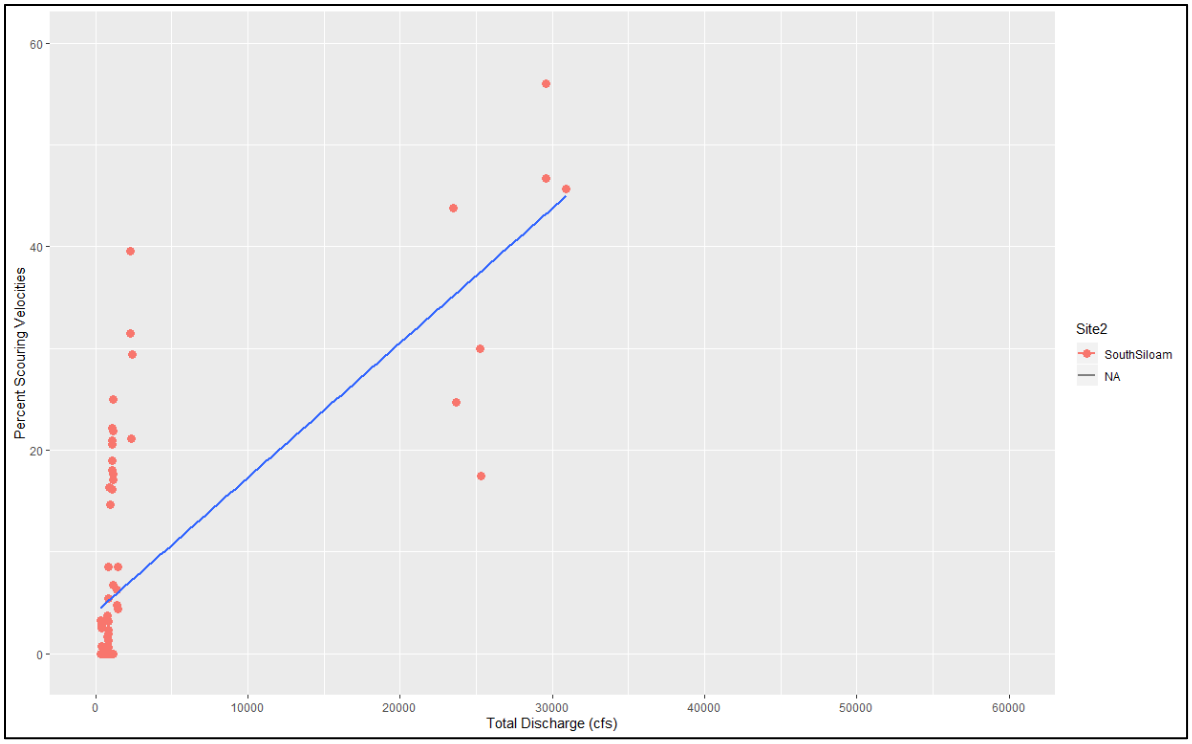


Figure 47 Illinois River at south Siloam Springs, relationship between percentage of scouring velocities and total daily average discharge.

Conclusions

The results for this scour analysis were aligned with results from other studies. Scour can be variable across similar flows and the location of scour within the channel was also often inconsistent. As the intensity of a scour event decreases, characteristics such as channel morphology and substrate have an increasing role in scour occurring or not occurring (Stevenson et al., 1996, Francouer and Biggs, 2006, Katz et al., 2018). Without a consistent clear pattern of scour events that were relatable to an accessible metric such as daily average flow this analysis does not readily assist in crafting an operational definition for the critical condition term. Additionally, this analysis was labor and expertise intensive and it is not realistic for it to be done repeatedly for each sampling event to determine if critical conditions were satisfied post hoc. Based on this outcome, OWRB staff relied upon the hydrograph separation analysis to translate the critical condition terminology into an operational definition.

Critical Condition Definition

The critical condition term was newly introduced by the Joint Study Committee and works to identify water quality samples utilized to evaluate the total phosphorus criterion. The analysis presented above served to translate the committee critical condition terminology into an operational definition that could be feasibly and consistently implemented in a myriad of water quality management programs. Based on this analysis the critical condition is defined as:

> *The critical condition is when baseflow is fifty-five (55%) or greater of the total daily average flow calculated by the USGS hydrograph separation method sliding-interval.*

This language is recommended for adoption into the Implementation of Oklahoma Water Quality Standards in section 785:46-15-14. This language will apply to data used in the Use Support Assessment Protocol for the Aesthetic beneficial use applicable to the scenic rivers reaches of Illinois River, Barren Fork Creek, and Flint Creek.

Additional Changes to Use Support Assessment Protocol for the Aesthetic Beneficial Use

Oklahoma's Use Support Assessment Protocols (USAP) are located in Implementation of Oklahoma Water Quality Standards 785:46, Subchapter 15. The USAP provides a consistent approach to analyzing water quality data for the assessment of beneficial uses and addresses necessary data requirements such as, minimum required samples. The revised TP criterion duration is a 6-month rolling average and the USAP prescribes the minimum number of samples to be included in each 6-month average calculation and the minimum number of 6-month averages needed to conduct a water quality assessment.

Calculation of 6-month Average

The OWRB Beneficial Use Monitoring Program (BUMP) conducts statewide ambient water quality monitoring, including at four sites in the Illinois River watershed. ADEQ has a similar water quality monitoring program and routinely monitors the location south of Siloam Springs at AR Highway 59, among other sites in the Arkansas portion of the watershed. Because the new critical condition term will restrict the total phosphorus data that can be used for water quality assessment a review was done to evaluate the availability of data generated by current monitoring programs that would meet the new restriction.

The recommended operational definition for the critical condition is when baseflow is fifty-five (55%) or greater of the total daily average flow. The entire record of sampling events from OWRB, USGS, and ADEQ within the Illinois River watershed from 2008-2018 were screened according to this critical condition definition; meaning that sampling events that occurred when baseflow was less than 55% were removed and the final population of sampling events were only those that satisfied the critical condition term. Results are presented below for the Tahlequah, Watts, and south Siloam Springs monitoring locations. At all three locations ~ 60 – 70% of months have at least one sample available to contribute to a 6-month average calculation (Figures 48). Additionally, when months where more than one sample was available the overall percentage of months with available samples increases to greater than 80% for all sites.

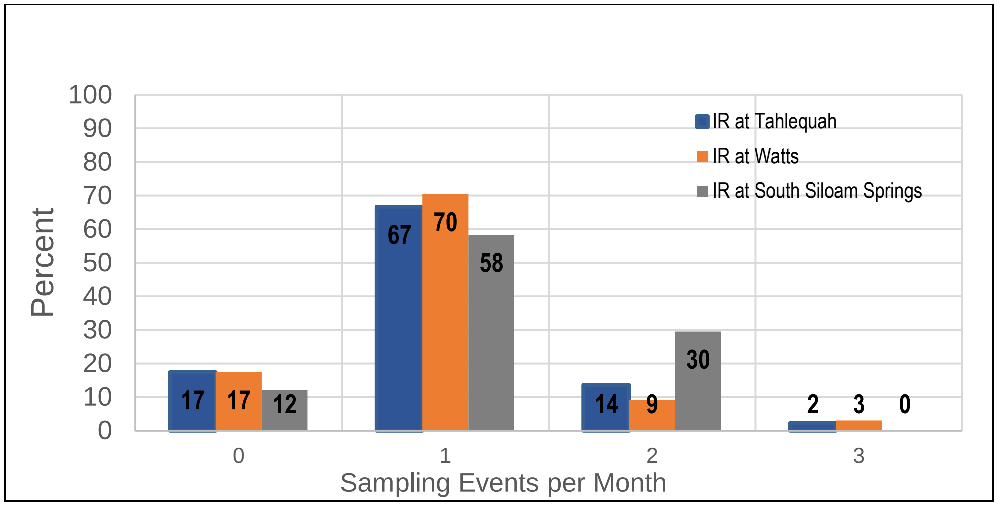


Figure 48 Percent of months with 0, 1, 2, or 3 samples eligible for inclusion in 6-month average calculation.

After reviewing the availability of data, two additional considerations were included. First, monitoring programs need some flexibility; it cannot be guaranteed that an eligible sample will be collected every month year in and year out. Second, it was desired to maximize the number of 6-month averages calculated and not sacrifice the calculation for any given month because the underlying requirement for individual monthly data values was overly stringent. As the minimum requirement for individual monthly data values increased the possibility of sacrificing a given month's 6-month average also increased. Thus, it was found that at least at least 4 individual measured data values were required for the calculation of 6-month average total phosphorus. This requirement is recommended for adoption in the USAP for the Aesthetic beneficial use applicable to the scenic rivers reaches of Illinois River, Barren Fork Creek, and Flint Creek (785:46-15-14).

Water Quality Assessment Requirements

The USAP typically requires at least ten samples for a water quality assessment for most parameters and beneficial uses (785:46-15-3). This same minimum "sample" number was carried forward for the revised TP criterion. A minimum of ten (10) calculated 6-month averages are required to assess the beneficial use within a one-year period. For the assessment of the five-year period, a minimum of thirty (30) 6-month averages are required. This ensures that there is a robust data set for the majority of years within the five-year period. The one-year and five-year assessment periods are the direct implementation of the criterion's two-part frequency.

References

Anderson, C.W. and K.D. Carpenter. 1998. Water Quality and Algal Conditions in the North Umpqua River Basin, Oregon, 1992-1995, and Implications for Resource Management. U.S. Department of the Interior, U.S. Geological Survey. Water Investigations Report 98-4125. Portland, OR. 78pp.

Arkansas TAG Members' Minority Report to OWRB, March 9, 2012. Retrieved from: www.owrb.ok.gov/quality/standards/scenicrivers.php

Barlow, P.M., Cunningham, W.L., Zhai, Tong, and Gray, Mark, 2015, U.S. Geological Survey Groundwater Toolbox, a graphical and mapping interface for analysis of hydrologic data (version 1.0) User guide for estimation of base flow, runoff, and groundwater recharge from streamflow data: U.S. Geological Survey Techniques and Methods, book 3, chapter B10

Biggs, B. J. 2000. Eutrophication of streams and rivers: Dissolved Nutrient-Chlorophyll

Busse, L.B., J. C. Simpson, S.D. Cooper. 2006. Relationship Among Nutrients, Algae, And Land Use in Urbanized Southern California Streams. Can. J. Aquat. Sci. 63: 2621-2638.

Chinnasamy,P. and Hubbart, J. 2015. Stream and Shallow groundwater nutrient concentrations in an Ozark forested riparian zone of the central USA, Environmental Earth Sciences 73, 6577-6590

Christensen, J.L., and Herrick, L.E., 1982, Mississippi River test, volume 1. Final report DCP4400/300, prepared for the U.S. Geological Survey by AMETEK/Straza Division, El Cajon, California, under contract No. 14–08–001–19003, p. A5–A10.

Clark, G., Mueller, D., Mast, M.A. 2000. Nutrient Concentrations and Yield in Undeveloped Stream Basins of the United States, Journal of the American Water Resources Association, Vol. 36, No. 4 849-860

Dodds, W. and Welch, E. 2000. Establishing Nutrient Criteria in Streams. J. N. Am. Benthol. Soc. 19(1): 186-196

Dodds, W., Gudder, D., 1992. The ecology of Cladophora, Journal of Phycology 28, 415-427

Fetscher, A., E., Stancheva, R., Kociolek, J., P., Sheath, R., Stein, E., Mazor, R., Ode, P., Busse, L. 2014. Development and comparison of stream indices of biotic integrity using diatoms vs. non-diatom algae vs a combination, Journal of Applied Phycology 26, 433-450

Final Report to Governors from the Joint Study Committee and Scientific Professional (2016, December 19). Principal Investigator: Ryan King, PhD. Retrieved from https://www.ok.gov/conservation/Agency_Divisions/Water_Quality_Division/IR_Joint_Study_Committee_Reference_Info.html

Flinders, C., Hart, D., 2009. Effects of pulsed flows on nuisance periphyton growths in rivers: a mesocosms study, River Research and Applications 25, 1320-1330

Fox, G., Purvis, R., Penn, C., 2016. Streambanks: A net source of sediment and phosphorus to streams and rivers, Journal of Environmental Management 181, 602-614

Francouer, S., Biggs, B., 2006. Short-term effects of elevated velocity and sediment abrasion on benthic algal communities, Hydrobioloigia 561, 59-69
Gonzales, A.L., Nonner, J., Heijkers, J., Uhlenbrook, S. 2009. Comparison of different base flow separation methods in a lowland catchment, Hydrology and Earth System Sciences, 13, 2055-2068

Granato G.E., Ries, K.G.III and Steeves, P.A., 2017 Compilation of streamflow statistics calculated from daily mean streamflow data collected during water years 1901-2015 for selected

U.S. Geological Survey stream gages: U.S. Geological Survey Open File Report 2017-1108

Katz, S., Segura, C., Warren D. 2018. The influence of channel bed disturbance on benthic Chlorophyll *a*: A high resolution perspective, Geomorphology, 305, 141-153

Kramer, D. 1987. Dissolved oxygen and fish behavior, Environmental Biology of Fishes, Vol. 18, No. 2 pp. 81-92

Lewis, JM., and Esralew, RA., 2009 Statistical Summaries of Streamflow in and near Oklahoma Through 2001: U.S. Geological Survey Scientific Investigations Report SIR 2009-5135

Linsley, R., Kohler, M., & Paulhus, J. (1975) Hydrology for Engineers (2nd ed.). McGraw-Hill

Macrae, M.,Zhang, Z., Stone,M., Price,J., Bourbonniere, R., Leach, M. 2011 Subsurface mobilization of phosphorus in an agricultural riparian zone in response to flooding from an upstream reservoir, Canadian Water Resources Journal, 36(4) 293-311

McCarty, James., Haggard, Brian., 2016 Can We Manage Nonpoint Source Pollution Using

Nutrient Concentrations during Seasonal Baseflow? Agricultural Environmental Letters, Vol. 1, Issue 1, 2471:9625

Miller, R., Robinson, H. (2004) Fishes of Oklahoma. University of Oklahoma Press.

Mueller, D.S., Wagner, C.R., Rehmel, M.S., Oberg, K.A,, and Rainville, Francois, 2013, Measuring discharge with acoustic Doppler current profilers from a moving boat (ver. 2.0, December 2013): U.S. Geological Survey Techniques and Methods, book 3, chap. A22, 95 p., http://dx.doi.org/10.3133/tm3A22.

Oberg, K.A., and Mueller, D.S., 2007b, Validation of streamflow measurements made with acoustic Doppler current profilers: Journal of Hydraulic Engineering, v. 133, no. 12, p. 1421–1432.

Oklahoma Administrative Code (OAC) Title 785, Chapter 45. Oklahoma's Water Quality Standards

Oklahoma Administrative Code (OAC) Title 785, Chapter 46. Implementation of Oklahoma's Water Quality Standards

Oklahoma Scenic Rivers Joint Phosphorus Study: Interim Report, 8 April 2016. Principal Investigator: Ryan King, PhD. Retrieved from https://www.ok.gov/conservation/Agency_Divisions/Water_Quality_Division/IR_Joint_Study_Committee_Reference_Info.html

Oklahoma Water Resources Board, March 30, 2012. Final Report, Oklahoma Scenic Rivers Phosphorus Criteria Review. Retrieved from: www.owrb.ok.gov/quality/standards/scenicrivers.php

Oklahoma Water Resources Board, May 2011. Quality Assurance Project Plan for Scenic Rivers Phosphorus Criterion Review. Retrieved from: www.owrb.ok.gov/quality/standards/scenicrivers.php

Pelletier, A., Andréassian, V., Hydrograph separation: an impartial parameterization for an imperfect method, Hydrology and Earth System Sciences, 2019-503

Rantz, S.E., and others, 1982, Measurement and computation of streamflow, volume 1, Measurement of discharge: U.S. Geological Survey Water-Supply Paper 2175, 631 p.

Relationships for Benthic Algae. J. N. Am. Benthol. Soc. 19:17-31.

Simpson, M.R., and Oltmann, R.N., 1993, Discharge measurement using an acoustic Doppler current profiler: U.S. Geological Survey Water-Supply Paper 2395, 34 p.

Sloto, Ronald., Crouse, Michele., 1996 HYSEP: A Computer Program for Streamflow

Hydrograph Separation and Analysis. U.S. Geological Survey, Water-Resources Investigations Report 96-4040

Stevenson, R.J., Bothwell, M., Lowe, R. (1996) Algal Ecology Freshwater Benthic Ecosystems. Academic Press Inc.

Turner, J., and Farley, T. 1971. Effects of Temperature, Salinity, and Dissolved Oxygen on the

Survival of Striped Bass Eggs, and Larvae, California Fish and Game 57(4): 268-273

U.S. Environmental Protection Agency, August 2017. *National Water Quality Inventory: Report to Congress*. EPA 841-R-16-011.

U.S. EPA, Nutrient Scientific Technical Exchange Partnership & Support (N-STEPS), Statistical Tools Fact Sheets, Change Point Analysis. Website last updated on October 23, 2020. https://www.epa.gov/nutrient-policy-data/n-steps-statistical-tools-fact-sheets Accessed on November 5, 2020

United States Department of Agriculture, United States Forest Service, 2003. A Soil Bioengineering Guide for Streambank and Lakeshore Stabilization, Chapter 2 – The Riparian Ecosystem. Accessed September 2020 https://www.fs.fed.us/t-d/php/library_card.php?p_num=FS-683P

US Army Corps of Engineers, Institute for Water Resources. 2019. *Value to the Nation Fast Facts: USACE Recreation 2019 Lake Report, Tenkiller Ferry Reservoir.* https://www.iwr.usace.army.mil/Missions/Value-to-the-Nation/Fast-Facts/ Accessed on July 17, 2020

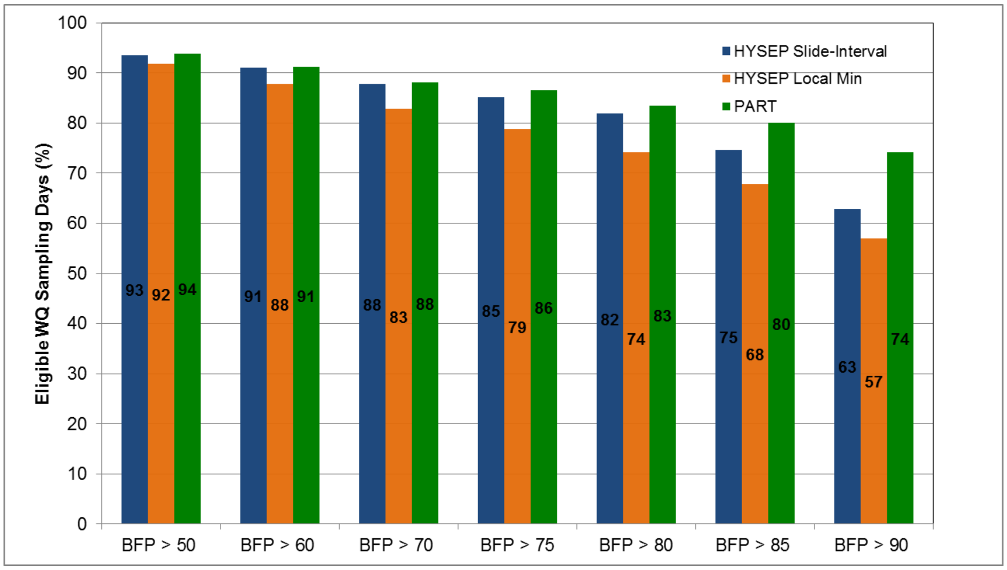
US EPA Causal Analysis/Diagnosis Decision Information System (CADDIS). Volume 2: Sources, Stressors & Responses. https://www.epa.gov/caddis Accessed on July 16, 2020

US EPA, Office of Water. Nutrient Criteria Technical Guidance Manual. EPA-822-B-00-002. July 2000

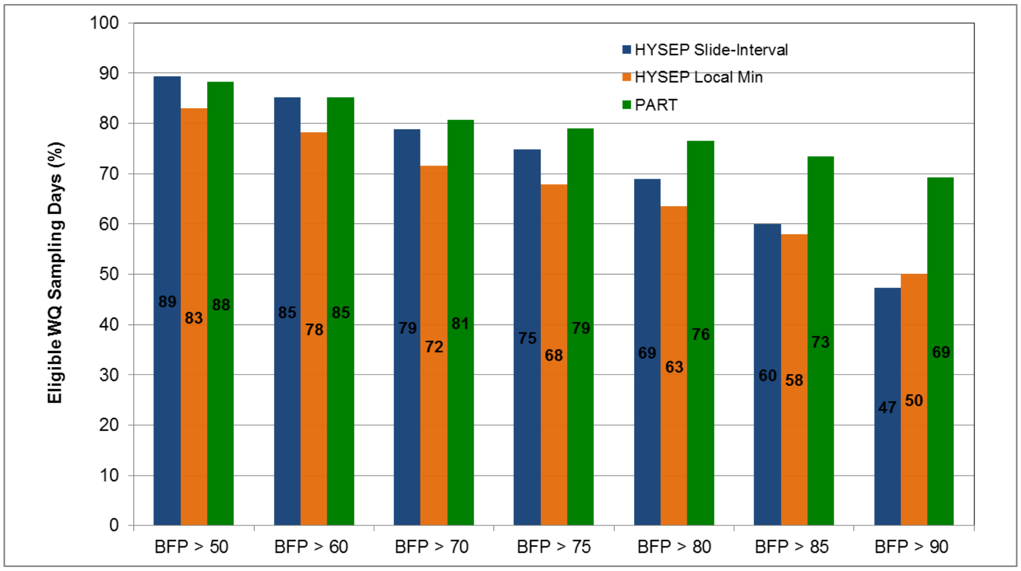
Welch, E.B. and J. M. Jacoby. 2004. Pollutant Effects in Freshwater: applied Limnology, Third edition. Spon Press. London, UK. 504 pp

Appendix 1

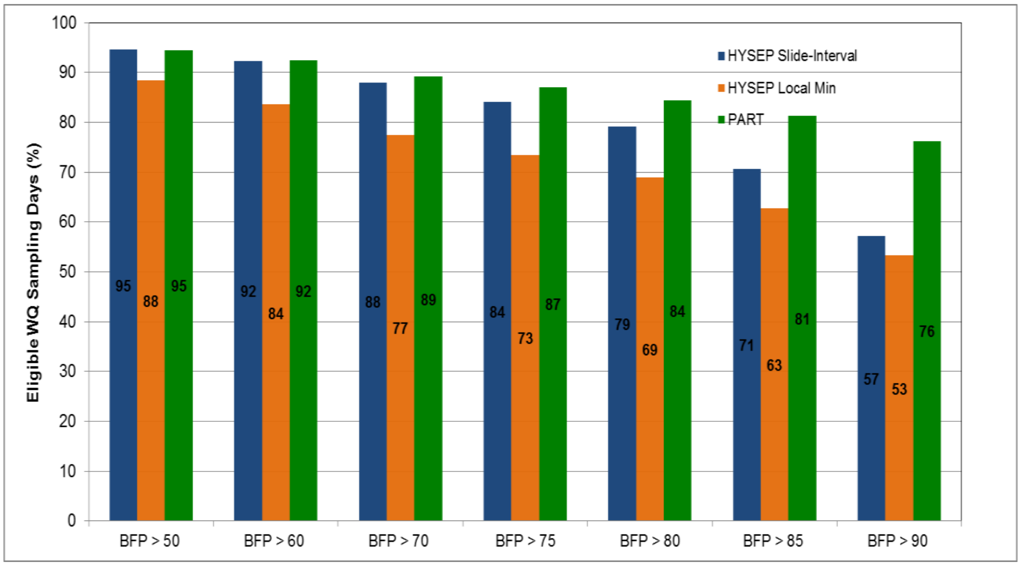
Hydrograph separation analysis results



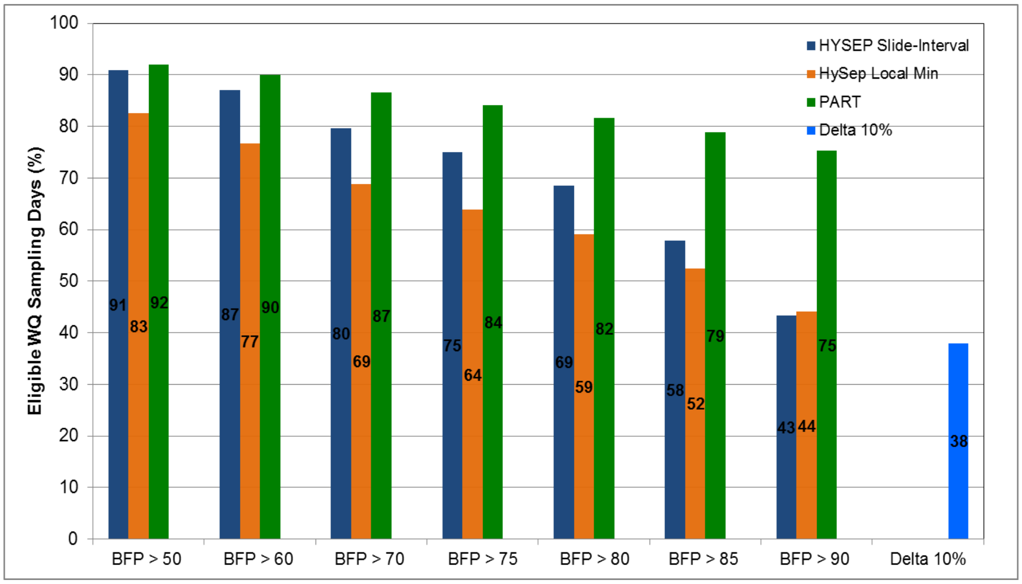
Osage Creek near Elm Springs



Illinois River at Savoy

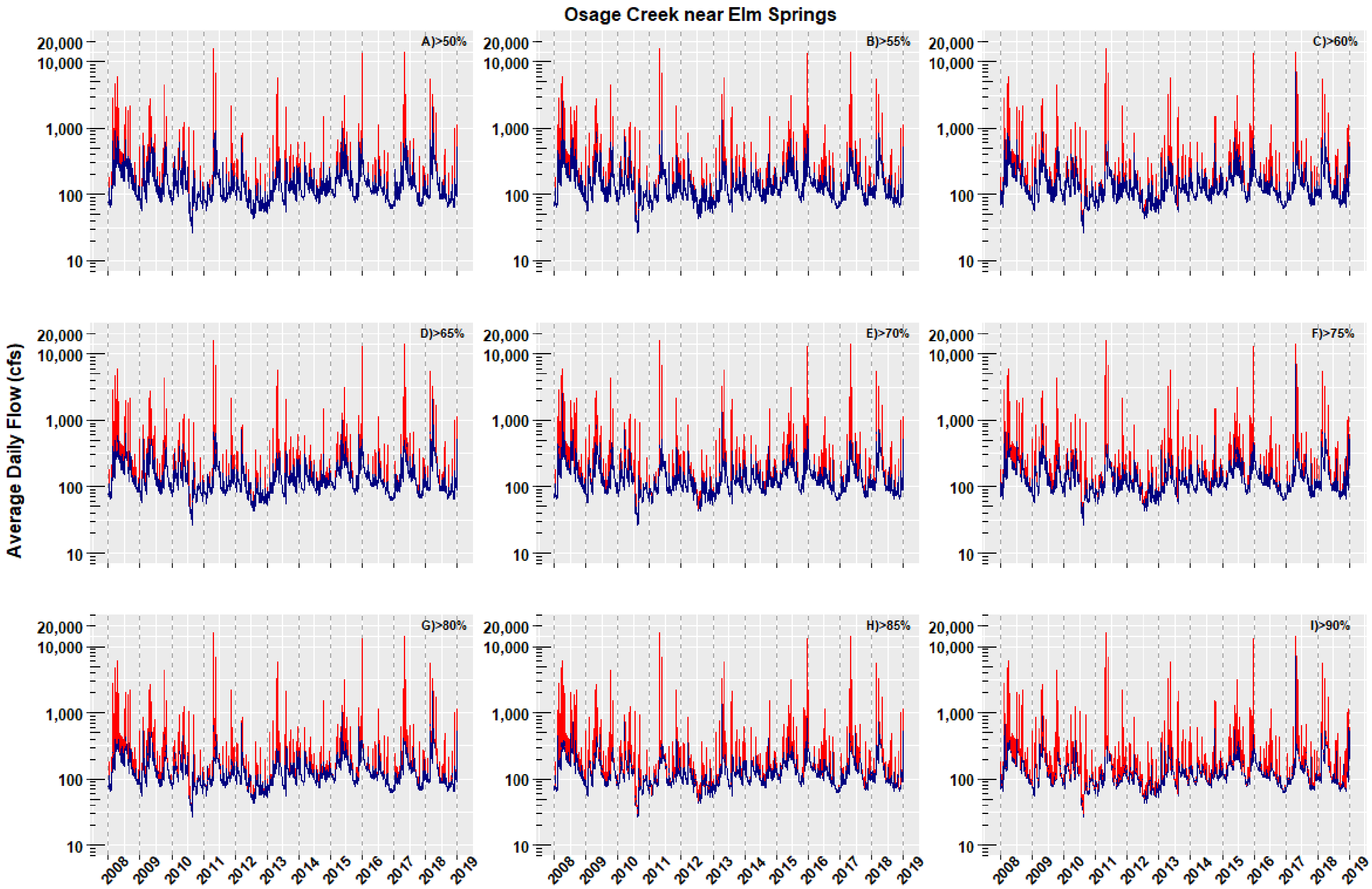


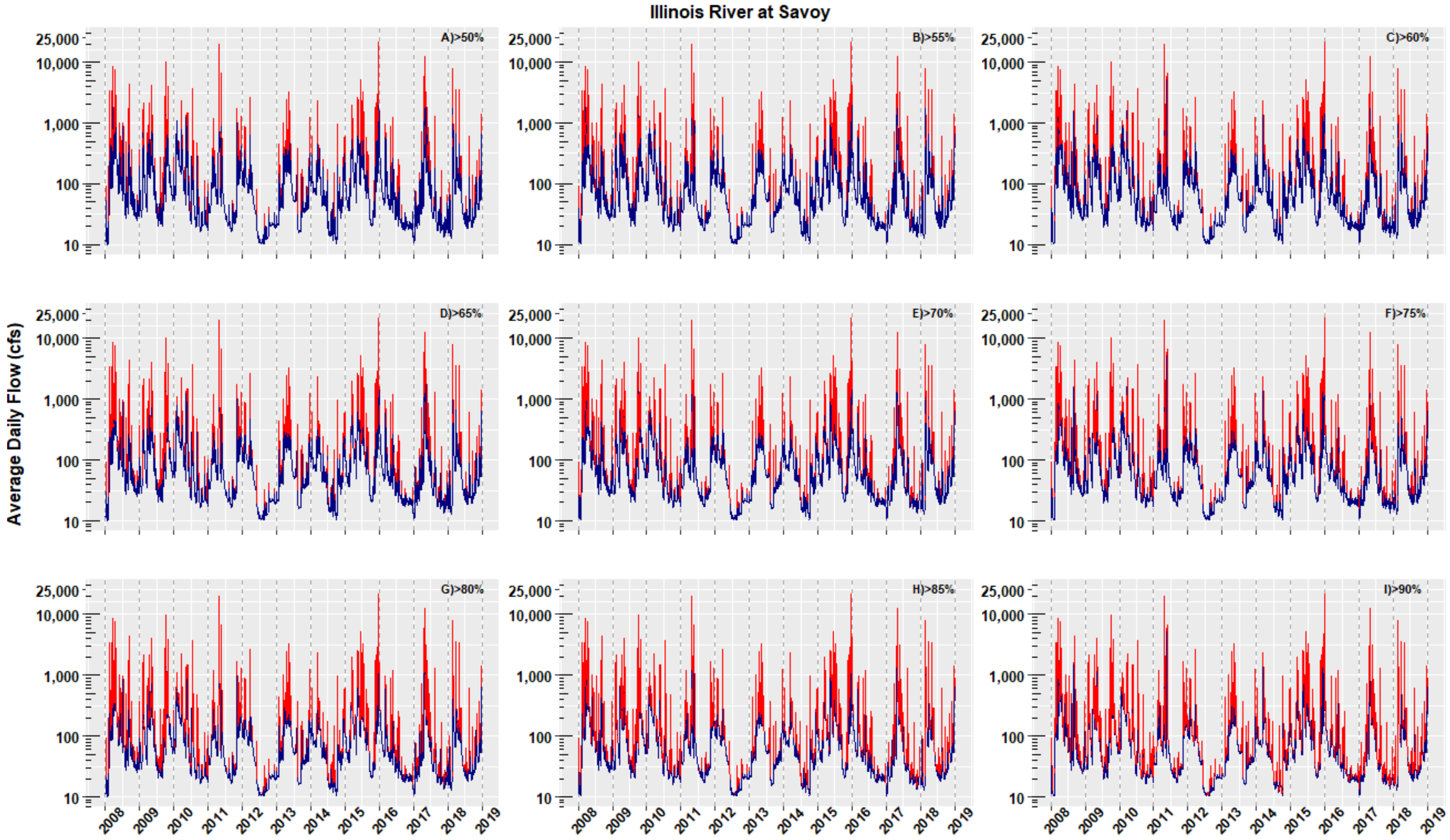
Flint Creek near Kansas

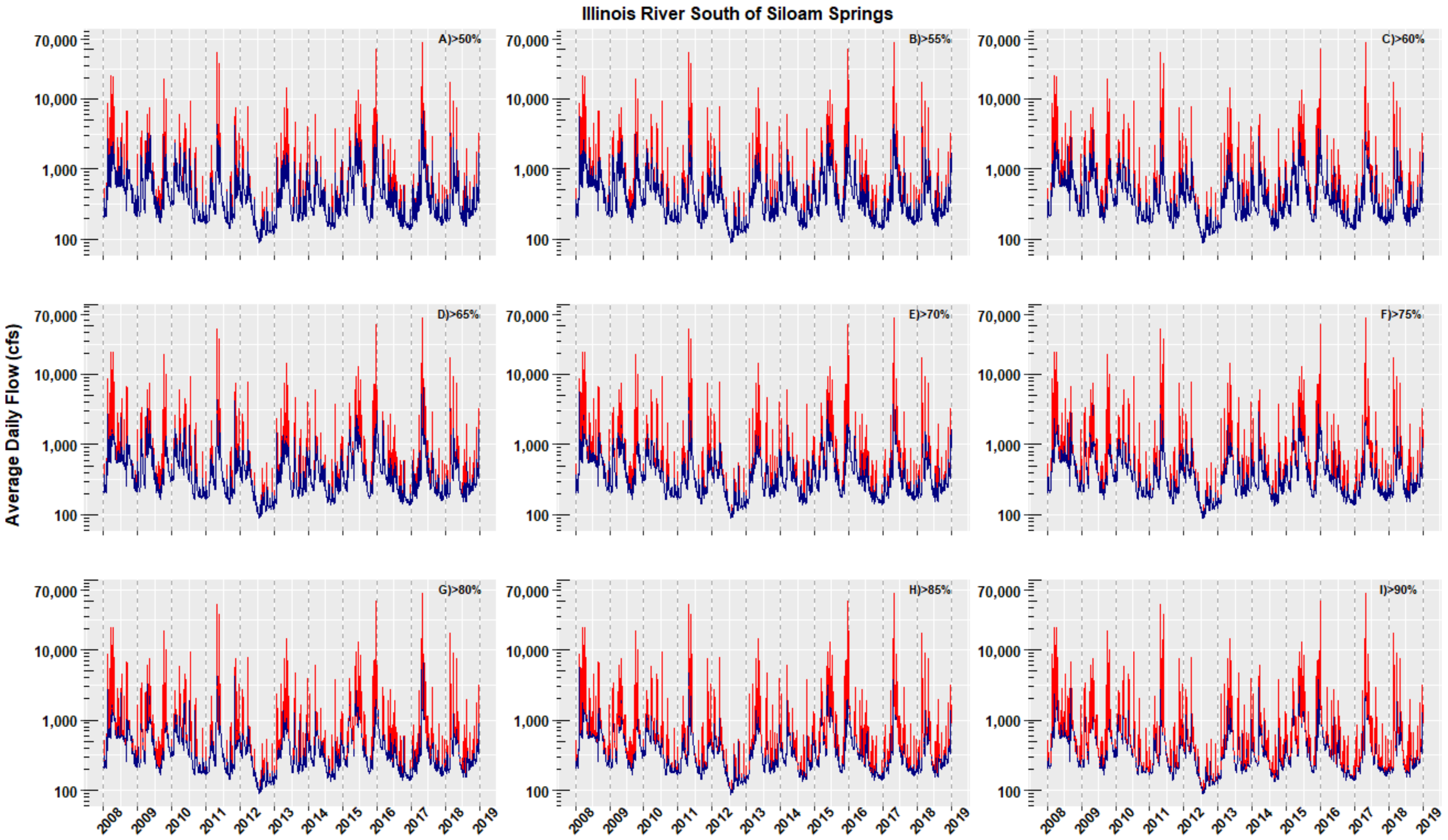


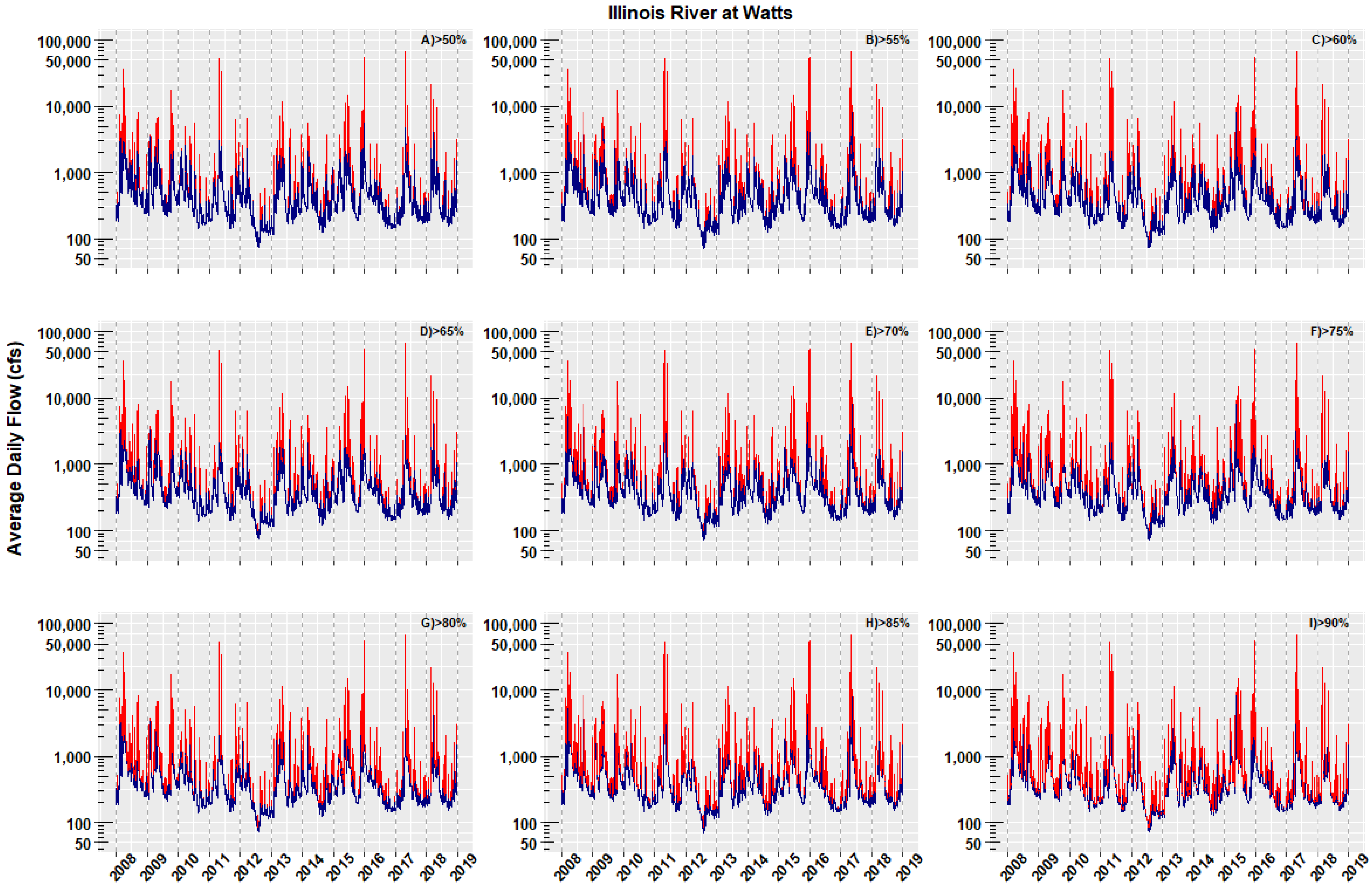
Barren Fork Creek at Eldon

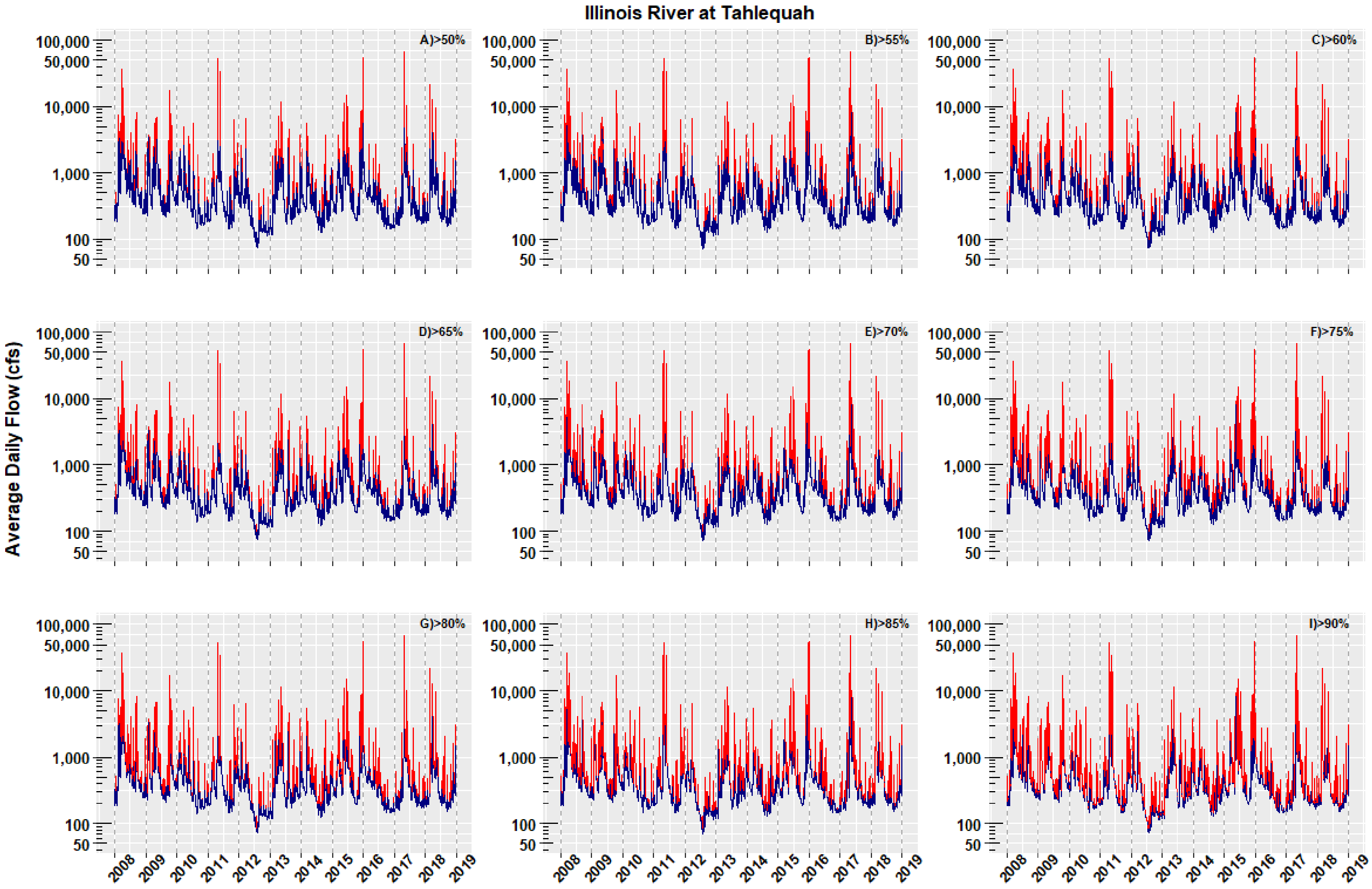
Hydrographs (Jan. 2008- Dec. 2018), red indicates portion of hydrograph ineligible for sampling and blue represents portion eligible for sampling, based on HYSEP sliding-interval method and baseflow percentage thresholds of > 50% through > 90%.

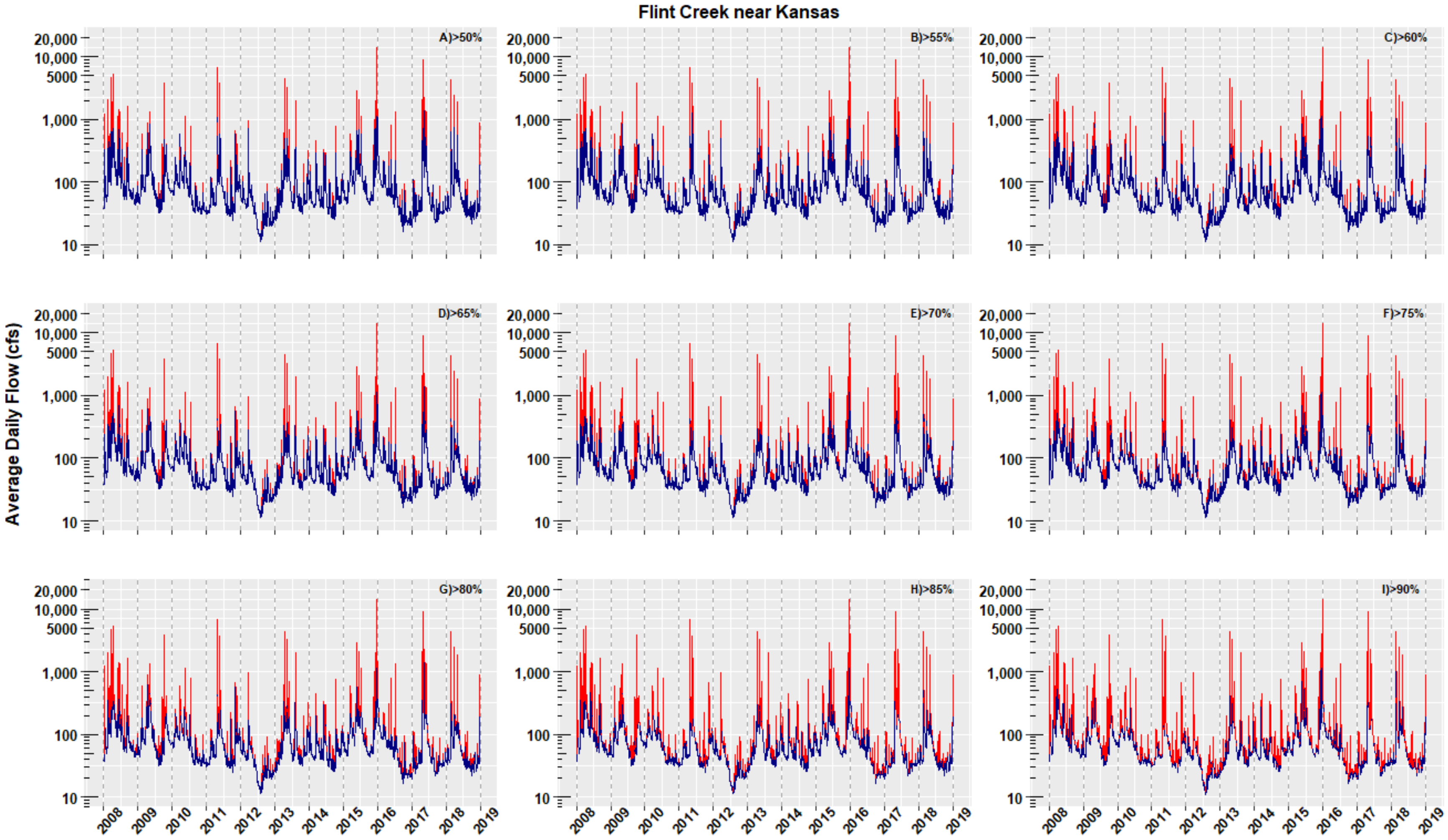

Osage Creek near Elm Springs
Average Daily Flow (cfs)
A)>50%
B)>55%
C)>60%
D)>65%
E)>70%
F)>75%
G)>80%
H)>85%
I)>90%
20,000
10,000
1,000
100
10
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019


Illinois River at Savoy
Average Daily Flow (cfs)
A)>50%
B)>55%
C)>60%
D)>65%
E)>70%
F)>75%
G)>80%
H)>85%
I)>90%
25,000
10,000
1,000
100
10
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019


Illinois River South of Siloam Springs
Average Daily Flow (cfs)
A)>50%
B)>55%
C)>60%
D)>65%
E)>70%
F)>75%
G)>80%
H)>85%
I)>90%
70,000
10,000
1,000
100
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019


Illinois River at Watts
Average Daily Flow (cfs)
A)>50%
B)>55%
C)>60%
D)>65%
E)>70%
F)>75%
G)>80%
H)>85%
I)>90%
100,000
50,000
10,000
1,000
100
50
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019


Illinois River at Tahlequah
Average Daily Flow (cfs)
A)>50%
B)>55%
C)>60%
D)>65%
E)>70%
F)>75%
G)>80%
H)>85%
I)>90%
100,000
50,000
10,000
1,000
100
50
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019


Flint Creek near Kansas
Average Daily Flow (cfs)
A)>50%
B)>55%
C)>60%
D)>65%
E)>70%
F)>75%
G)>80%
H)>85%
I)>90%
20,000
10,000
5000
1,000
100
10
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019

Baron Fork Creek near Eldon