Montana Department of
Environmental Quality

Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana’s Wadeable Streams and Rivers:

Update 1

May 2013

Prepared by:
Water Quality Planning Bureau, Standards Section
Montana Department of Environmental Quality
Water Quality Planning Bureau
1520 E. Sixth Avenue
P.O. Box 200901
Helena, MT 59620-0901

WQPBWQSTR-002

Suggested citation: Suplee, M.W[1]., and V. Watson[2], 2013. Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers—Update 1. Helena, MT: Montana Dept. of Environmental Quality.

---

[1] Environmental Science Specialist, Montana Department of Environmental Quality, 1520 East 6th Ave, Helena, MT, 59620, msuplee@mt.gov;

[2] Professor, Environmental Studies Program, University of Montana, Missoula, MT, 59812.

EXECUTIVE SUMMARY

This document is the first update by the Department of Environmental Quality (Department) to the numeric nutrient criteria recommendations it made in 2008. The science of eutrophication in general and numeric nutrient criteria in particular has continued to advance in the interim years. In addition, the Department has modified somewhat the process by which numeric criteria are derived. In 2008, the Department used ecoregions, stressor-response studies (nutrient as stressor, impact to stream beneficial use as response), and data from reference streams to develop criteria. Various cases studies had established a linkage between nutrient concentrations observed in reference streams and harm to beneficial uses; on average, harm-to-use occurred at the equivalent of the 86th percentile of the nutrient reference distribution. In the 2008 document the Department relied heavily on two percentiles from the ecoregional reference distributions, namely the 75th and 90th, to derive criteria for each ecoregion.

The approach taken in 2008 had its shortcomings, however. In some ecoregions the method resulted in criterion concentrations which other data and studies have shown were unnecessarily stringent, while in other ecoregions the method resulted in criteria at concentrations that were too high (not protective). The method— albeit simple, consistent, and transparent—limited the Department's ability to derive best-fit criteria for each ecoregion. Fundamentally, the Department considers the combined use of ecoregions, stressor-response studies, and reference data to be a sound approach (as did the external peer reviewers of the draft of this document). But compared to 2008, more stressor-response studies are now available and these can better inform the criteria derivation process. As a result, in this update there has been less reliance on specific reference-distribution percentiles and much more reliance on regional as well as non-regional stressor-response studies.

The other major change, relative to the 2008 document, is that the Department will not be recommending nitrate (or nitrate + nitrite) criteria for adoption for the control of eutrophication at this time. Only total phosphorus (TP) and total nitrogen (TN) criteria are provided here. Rapid uptake of soluble nitrogen compounds by aquatic organisms (mainly algae and plants) makes these compounds' concentrations quite variable, and difficult to use as ambient surface water criteria. Total nitrogen and TP provide better overall correlation to eutrophication response than soluble nutrients, and are generally more practical than soluble forms for ambient river monitoring and assessment, total maximum daily loads, etc.

The Department is recommending both TN and TP criteria for stream protection. Phosphorus (P) control is sometimes promoted as the only approach needed to limit eutrophication, this being based largely on the more economical removal of P from wastewater and the assumption that P can be made to become limiting in the waterbody. But data pertaining to streams and rivers indicate that it would be unwise to adopt only P criteria. Mixed assemblages of benthic algae are very often limited by nitrogen or nitrogen and phosphorus (co-limitation) in the region's flowing waters. A P-only approach, in order to work, would require that P standards be set to the very low background levels observed in our western region's reference sites (e.g., 10 µg TP/L). If the P standard were not set to natural background, and no controls on N were undertaken, then the commonly occurring N limitation or N and P co-limitation would lead to algal growth stimulation nonetheless. Worse yet, in the long term, a P-only strategy would result in highly skewed (elevated) N:P ratios accompanying the low P levels. These management-induced conditions might control green algae biomass but may lead to nuisance blooms of the diatom algae *Didymosphenia geminata,* which has in recent years formed nuisance blooms in rivers and streams in Montana and word-wide.

A balanced and prudent policy would be to reduce both N and P and maintain, as nutrient concentration reductions occur, a roughly balanced (i.e., Redfield) ratio between the two. This is the strategy that has been applied on the Clark Fork River (where nutrients standards were adopted in 2002) and it appears to be working there. In addition, other researchers in the field are recommending that both N and P need to be controlled to effectively manage eutrophication. Thus, both N and P criteria for wadeable streams and rivers are proposed in this document.

The document has been organized so that readers can quickly locate key information pertaining to an ecoregion of interest. Data and discussion specific to each ecoregion are then presented on three to four pages. A map of Montana showing the ecoregion in which the criteria apply is shown first, followed by the criteria recommendations, and then tables of descriptive statistics for the ecoregion's reference streams. Then readers will find TN and TP histograms for the reference data, a discussion of the scientific studies (regional and beyond) that were used to help derive the criteria, other considerations pertaining to the derivation of the criteria, and a conclusion with final thoughts about the criteria.

The Department recognizes that within each ecoregional zone there are likely to be some streams with unique characteristics that could render the ecoregional criteria inappropriate. These characteristics include, for example, the presence of a large dam-regulated lake or reservoir upstream, or the upstream influence of a level-IV ecoregion known to have naturally elevated TP concentrations. A few cases have already been identified, and reach-specific criteria for them are presented and discussed in the document.

Below are summarized the criteria concentrations that have been recommended (**Table ES-1**). As was the case in the 2008 document, the criteria should generally apply seasonally.

Table ES-1. Recommended Numeric Nutrient Criteria for Different Montana Ecoregions and Stream Reaches.
Related assessment information is also shown.

		Parameter		
Ecoregion (level III or IV) and number, or Reach Description	**Period When Criteria Apply**	**Total Phosphorus (µg/L)**	**Total Nitrogen (µg/L)**	**Related Assessment Information***
Northern Rockies (15)	July 1 to September 30	25	275	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Canadian Rockies (41)	July 1 to September 30	25	325	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Idaho Batholith (16)	July 1 to September 30	25	275	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Middle Rockies (17)	July 1 to September 30	30	300	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Absaroka-Gallatin Volcanic Mountains (17i)	July 1 to September 30	105	250	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Northwestern Glaciated Plains (42)	June 16 to September 30	110	1300	
Sweetgrass Upland (42l), Milk River Pothole Upland (42n), Rocky Mountain Front Foothill Potholes (42q), and Foothill Grassland (42r)	July 1 to September 30	80	560	165 mg Chl*a*/m^2 and 70 g AFDM/m^2

Table ES-1. Recommended Numeric Nutrient Criteria for Different Montana Ecoregions and Stream Reaches.
Related assessment information is also shown.

		Parameter		
Ecoregion (level III or IV) and number, or Reach Description	**Period When Criteria Apply**	**Total Phosphorus (µg/L)**	**Total Nitrogen (µg/L)**	**Related Assessment Information***
Northwestern Great Plains (43) and Wyoming Basin (18)	July 1 to September 30	150	1300	
River Breaks (43c)	NONE RECOMMENDED	NONE RECOMMENDED	NONE RECOMMENDED	
Non-calcareous Foothill Grassland (43s), Shields-Smith Valleys (43t), Limy Foothill Grassland (43u), Pryor-Bighorn Foothills (43v), and Unglaciated Montana High Plains (43o)†	July 1 to September 30	33	440	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
INDIVIDUAL REACHES:				
Flint Creek, from Georgetown Lake outlet to the ecoregion 17ak boundary (46.4002, -113.3055)	July 1 to September 30	72	500	150 mg Chl*a*/m^2 and 45 g AFDM/m^2
Bozeman Creek, from headwaters to Forest Service Boundary (45.5833, -111.0184)	July 1 to September 30	105	250	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Bozeman Creek, from Forest Service Boundary (45.5833, -111.0184) to mouth at East Gallatin River	July 1 to September 30	76	270	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Hyalite Creek, from headwaters to Forest Service Boundary (45.5833,-111.0835)	July 1 to September 30	105	250	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Hyalite Creek, from Forest Service Boundary (45.5833,-111.0835) to mouth at East Gallatin River	July 1 to September 30	90	260	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
East Gallatin River between Bozeman Creek and Bridger Creek confluences	July 1 to September 30	50	290	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
East Gallatin River between Bridger Creek and Hyalite Creek confluences	July 1 to September 30	40	300	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
East Gallatin River between Hyalite Creek and Smith Creek confluences	July 1 to September 30	60	290	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
East Gallatin River from Smith Creek confluence to the mouth (Gallatin River)	July 1 to September 30	40	300	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
*Benthic algae density.				

† For the Unglaciated High Plains ecoregion (43o), criteria only apply to the polygon located just south of Great Falls, MT.

TABLE OF CONTENTS

Acronyms xi

1.0 Introduction 1-1

2.0 Methods Used to Derive the Criteria 2-1

2.1 Ecoregions as the Basis for Establishing Nutrient Criteria Zones 2-1

2.2 Criteria in this Document Apply to Wadeable Streams 2-1

2.3 Criteria Apply Seasonally, With Exceptions 2-2

2.4 Method for Assuring that All Reference Sites are Equitably Represented in an Ecoregion (The All-observations Dataset after applying Brillouin Evenness Index, And the Median dataset) 2-2

2.5 Literature Consulted 2-4

2.5.1 Literature Pertaining to Nutrient Enhancements in Rivers and Streams 2-4

2.6 Both Nitrogen and Phosphorus Criteria are Recommended 2-5

3.0 Ecoregion-specific Numeric Nutrient Criteria Recommendations 3-1

3.1 Level III: Middle Rockies (Ecoregion 17) 3-2

3.1.1 Level IV Ecoregion within the Middle Rockies: Absaroka-Gallatin Volcanic Mountains (17i) 3-5

3.2 Level III: Northern Rockies (Ecoregion 15) 3-8

3.3 Level III: Canadian Rockies (Ecoregion 41) 3-11

3.4 Level III: Idaho Batholith (Ecoregion 16) 3-14

3.5 Level III: Northwestern Glaciated Plains (Ecoregion 42) 3-17

3.5.1 Transitional Level IV Ecoregions within the Northwestern Glaciated Plains: Sweetgrass Upland (42l), Milk River Pothole Upland (42n), Rocky Mountain Front Foothill Potholes (42q), and Foothill Grassland (42r) 3-21

3.6 Level III: Northwestern Great Plains (Ecoregion 43) and the Wyoming Basin (Ecoregion 18) 3-25

3.6.1 Level IV Ecoregion within the Northwestern Great Plains: River Breaks (43c) 3-30

3.6.2 Transitional Level IV Ecoregions within the Northwestern Great Plains: Non-calcareous Foothill Grassland (43s), Shields-smith Valleys (43t), Limy Foothill Grassland (43u), Pryor-Bighorn Foothills (43v), and Parts of the Unglaciated Montana High Plains (43o) 3-34

4.0 Reach-specific Numeric Nutrient Criteria Recommendations 4-1

4.1 Flint Creek: From the Georgetown Lake Outlet to The Boundary of Ecoregion 17ak at Latitude 46.4002, Longitude -113.3055 4-1

4.2 Bozeman Creek, Hyalite Creek, And East Gallatin River 4-4

5.0 Acknowledgements 5-1

6.0 References 6-1

Appendix A – Peer-Review Comments and Responses A-1

LIST OF TABLES

Table ES-1. Recommended Numeric Nutrient Criteria for Different Montana Ecoregions and Stream Reaches. ii
Table 2-1. Large river segments within the state of Montana. 2-1
Table 2-2. Start and Ending Dates for Three Seasons (Winter, Runoff and Growing), by Level III Ecoregion. 2-2
Table 3-1A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Middle Rockies ecoregion. 3-2
Table 3-1B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Middle Rockies ecoregion. 3-3
Table 3-2A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Absaroka-Gallatin Volcanic Mountains (17i) ecoregion. 3-5
Table 3-2B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Absaroka-Gallatin Volcanic Mountains level-IV ecoregion. 3-5
Table 3-3A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northern Rockies ecoregion. 3-8
Table 3-3B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northern Rockies ecoregion. 3-8
Table 3-4A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Canadian Rockies ecoregion. 3-11
Table 3-4B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Canadian Rockies ecoregion. 3-11
Table 3-5A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Idaho Batholith ecoregion........ 3-14
Table 3-5B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Idaho Batholith ecoregion........ 3-14
Table 3-6A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northwestern Glaciated Plains ecoregion. 3-17
Table 3-6B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northwestern Glaciated Plains ecoregion. 3-17
Table 3-7A. Descriptive Statistics for TN and TP Concentrations in Transitional Level IV Ecoregions (42q, 42r) of the Northwestern Glaciated Plains. No data were available for 42l, 42n........ 3-21
Table 3-7B. Descriptive Statistics for TN and TP Concentrations in Transitional Level IV Ecoregions (42q, 42r) of the Northwestern Glaciated Plains. 3-22
Table 3-8. Site-average Benthic Algae Levels Measured in Streams of Ecoregions 42r and 42q. 3-23
Table 3-9A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northwestern Great Plains ecoregion. 3-25
Table 3-9B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northwestern Great Plains ecoregion. 3-25
Table 3-10A. Descriptive Statistics for TN and TP Concentrations in Reference Streams of the River Breaks (43c) level IV ecoregion. 3-30
Table 3-10B. Descriptive Statistics for TN and TP Concentrations in Reference Streams of the River Breaks (43c) level IV Ecoregion. 3-30
Table 3-10C. Descriptive Statistics for NO_{2+3} and SRP Concentrations in Reference Streams of the River Breaks (43c) level IV Ecoregion. 3-31

Table 3-11A. Descriptive Statistics for TN and TP Concentrations in Transitional Level IV Ecoregions (43s, 43t, 43u) of the Northwestern Great Plains. 3-34
Table 3-11B. Descriptive Statistics for TN and TP Concentrations in Transitional Level IV Ecoregions (43s, 43t, 43u) of the Northwestern Great Plains. 3-35
Table 4-1. Descriptive Statistics for TN and TP concentrations in Flint Creek just Below Georgetown Lake Outlet (July through September). 4-2
Table 4-2. Recommended Criteria for Reaches of Bozeman Creek, Hyalite Creek, and the East Gallatin River. 4-4
Table 4-3. Ecoregion-specific Reference Conditions and Numeric Nutrient Criteria for TN and TP. Data are from the all-observations dataset after applying the Brillouin Evenness Index....4-6
Table 4-4. Total Phosphorus Natural Background and Derived Nutrient Criteria for Example Stream and River Reaches in the East Gallatin River Watershed....4-7
Table 4-5. Total Nitrogen Natural Background and Derived Criteria for Example Stream and River Reaches in the East Gallatin River Watershed....4-7

LIST OF FIGURES

Figure 1-1. Overview of approach used to derive nutrient criteria in 2008 (Suplee et al., 2008).....1-2
Figure 1-2. Overview of approach used to derive nutrient criteria in this document....1-3
Figure 3-1. Map of Montana showing location of stream references sites (white dots). 3-1
Figure 3-2. Map of Montana showing the Middle Rockies ecoregion in gray....3-2
Figure 3-3. Nutrient concentrations from reference streams in the Middle Rockies ecoregion. 3-3
Figure 3-4. Map of Montana showing the Absaroka-Gallatin volcanic Mountains (17i), a level IV ecoregion within the Middle Rockies ecoregion. 3-5
Figure 3-5. Nutrient concentrations from reference streams in the Absaroka-Gallatin Volcanic Mountains (17i) ecoregion....3-6
Figure 3-6. Map of Montana showing Northern Rockies ecoregion. 3-8
Figure 3-7. Nutrient concentrations from reference streams in the Northern Rockies ecoregion....3-9
Figure 3-8. Map of Montana showing Canadian Rockies ecoregion. 3-11
Figure 3-9. Nutrient concentrations from reference streams in the Canadian Rockies ecoregion....3-12
Figure 3-10. Map of Montana showing the Idaho Batholith ecoregion. 3-14
Figure 3-11. Nutrient concentrations from reference streams in the Idaho Batholith ecoregion....3-15
Figure 3-12. Map of Montana showing the Northwestern Glaciated Plains ecoregion (42) in light gray......3-17
Figure 3-13. Nutrient concentrations from reference streams in the Northwestern Glaciated Plains ecoregion, but excluding data from the level IV ecoregions 42l, 42n, 42q, and 42r....3-18
Figure 3-14. Map of Montana showing in gray the transitional level IV ecoregions (42l, 42n, 42q, and 42r) within the Northwestern Glaciated Plains....3-21
Figure 3-15. Nutrient concentrations from reference streams in the Transitional Level IV ecoregions (42q, and 42r) of the Northwestern Glaciated Plains....3-22
Figure 3-16. Map of Montana showing the Northwestern Great Plains ecoregion (43) in gray, the Wyoming Basin (18) in light gray. 3-25
Figure 3-17. Nutrient concentrations from reference streams in the Northwestern Great Plains ecoregion (43), but excluding data from the mountain-to-plains transitional level IV ecoregions 43s, 43t, 43u, and 43v....3-26
Figure 3-18. Map of Montana showing the River Breaks (43c), a level IV ecoregion within the Northwestern Great Plains ecoregion. 3-30

Figure 3-19. Nutrient concentrations from reference streams in the River Breaks (43c) level IV ecoregion. 3-31
Figure 3-20. Benthic algal density (mg Chl*a*/m^2), all replicates, from reference streams in the River Breaks (43c) ecoregion......3-33
Figure 3-21. Map of Montana showing in gray the transitional level IV ecoregions (43s, 43t, 43u, and 43v) within the Northwestern Great Plains. 3-34
Figure 3-22. Nutrient concentrations from reference streams in the transitional level IV ecoregions (43s, 43t, 43u) of the Northwestern Great Plains. 3-35
Figure 3-23. Site-average benthic algae density (mg Chl*a*/m^2) from reference streams in the transitional level IV ecoregions (43s, 43t, 43u) of the Northwestern Great Plains ecoregion. 3-36
Figure 4-1. Map showing the Flint Creek watershed below the Georgetown Lake outlet......4-1
Figure 4-2. Nutrient concentrations observed in Flint Creek just downstream of the point where water exits Georgetown Lake through the dam. 4-2
Figure 4-3. Map showing the East Gallatin River watershed, including Bozeman Creek and Hyalite Creek. 4-4

ACRONYMS

Acronym	**Definition**
AFDM	Ash Free Dry Mass
ARM	Administrative Rules of Montana
Chl*a*	Chlorophyll-*a*
DEQ	Department of Environmental Quality (Montana)
DO	Dissolved Oxygen
HBI	Hilsenhoff Biotic Metric
NB	Natural Background
SOP	Standard Operating Procedure
SRP	Soluble Reactive Phosphate
TDP	Total Dissolved Phosphate
TKN	Total Kjeldahl Nitrogen
TN	Total Nitrogen
TP	Total Phosphorus
TSS	Total Suspended Solids
USGS	United States Geological Survey
WMA	Wildlife Management Area
WWTP	Wastewater Treatment Plant

1.0 INTRODUCTION

This is the first update to the Department of Environmental Quality (Department) document "Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers" (Suplee et al., 2008). Suplee et al. (2008) addresses methods that were used to derive numeric nutrient (nitrogen and phosphorus) criteria. The science of eutrophication in general and numeric nutrient criteria in particular has continued to advance in the interim years. Thus, this update reflects the most up-to-date nitrogen and phosphorus criteria recommendations for the control of eutrophication in wadeable streams and rivers that the Department has so far provided. With these revisions to the nutrient criteria, it bears repeating that the purpose of water quality criteria and standards is to define a level of a pollutant that will protect beneficial uses. This is the level to which degraded streams need to be restored; streams with water quality better than the criteria are addressed by the state's nondegradation provisions (i.e. ARM 17.30.701 through 17.30.718).

In the 2008 document, the Department used ecoregions (Woods et al., 2002), regional stressor-response studies (nutrient as stressor, impact to stream beneficial use as response), and data from reference streams to derive the criteria. Ecoregions were used to segregate the landscape into zones within which single nitrogen and phosphorus criteria—protective of the streams' beneficial uses and unique to each ecoregion—were recommended. Linkages had been made between harm to beneficial uses and nutrient concentrations which occurred, on average, at the 86th percentile of reference, with a coefficient of variation of ± 13% (i.e., from the 73rd to the 99th percentile; (Suplee et al., 2007). In developing its 2008 criteria recommendations, the Department relied heavily on two percentiles from the ecoregional reference distributions, namely the 75th and 90th (Suplee et al., 2008).

In 2008 the Department used only two different reference percentiles to derive the criteria because there were fewer regional dose-response studies available then. Further, the Department believed that it was best to be consistent in the use of reference-percentiles across broad areas of the landscape, because it would be fair and transparent. However in retrospect this approach had its failings, because in some ecoregions (e.g., the Canadian Rockies) the method resulted in criterion concentrations (6 µg TP/L) which other data and studies have shown to be unnecessarily stringent, while in other ecoregions (e.g., the Middle Rockies) the approach produced criteria concentrations (48 µg TP/L) we now believe to be somewhat too high. The original approach limited the Department's ability to recommend custom-fit criteria for different ecoregions that best reflect the level of water quality needed to protect the beneficial uses of each particular region's streams.

The Department still considers the combined use of ecoregions, stressor-response studies, and reference data to be a sound approach, but it was in need of modification. New stressor-response studies are now available and we believe these can better inform the criteria derivation process. In this update, which documents the Department's revised methods and recommended criteria, there will be less reliance on specified reference-distribution percentiles and much more reliance on regional as well as non-regional stressor-response studies. For clarity, we contrast below the approach taken in 2008 (**Figure 1-1**) vs. the approach taken in this document (**Figure 1-2**).

Another concern pertaining to the earlier work was the degree to which all reference sites within an ecoregion were equitably represented. In some ecoregions, a great deal of data has been collected at one or two reference sites and much less data at other sites. In this updated work, we improved objectivity by using two quantitative methods to assure that each reference site in an ecoregion only contributes a comparable amount of information to the ecoregional dataset; this is reflected in box 4 of **Figure 1-2** and is detailed in **Section 2.4**.

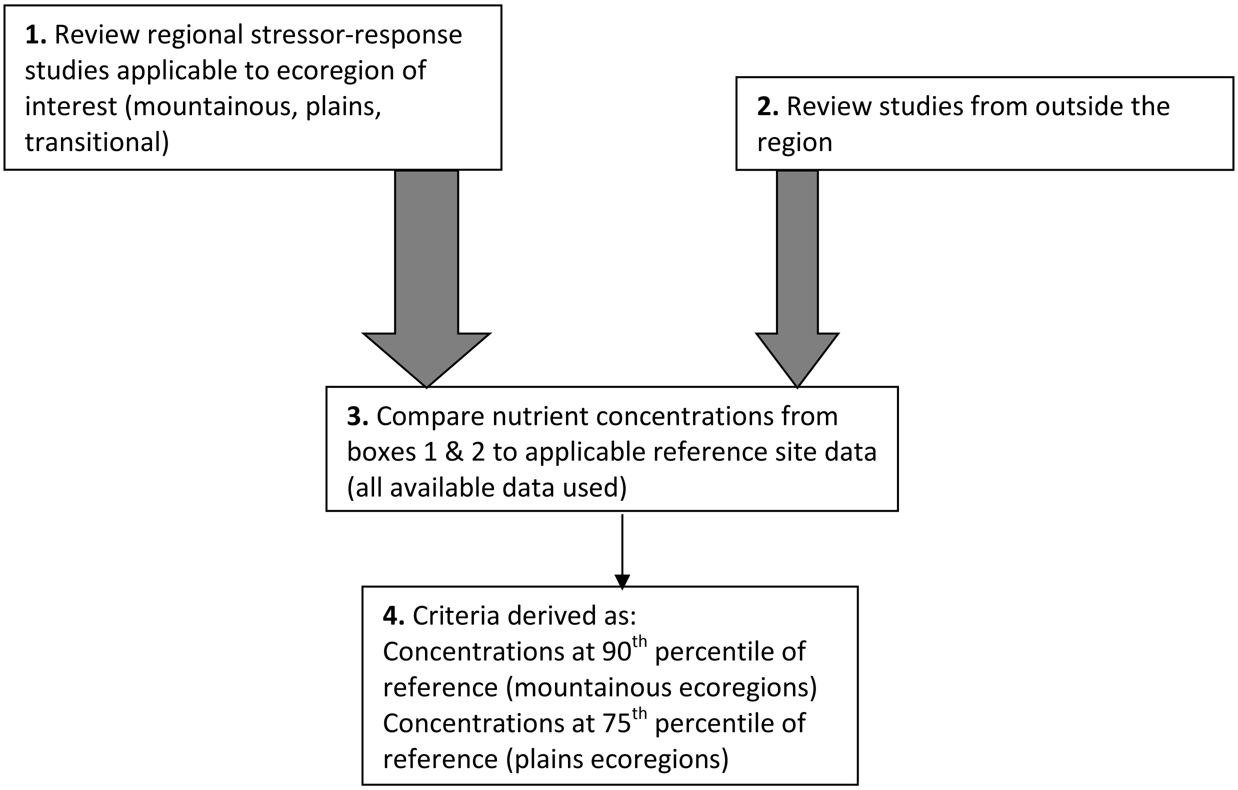

Figure 1-1. Overview of approach used to derive nutrient criteria in 2008 (Suplee et al., 2008). The size of the large grey arrows near the top of the figure represent the relative importance of the two information sources for deriving regional nutrient criteria.

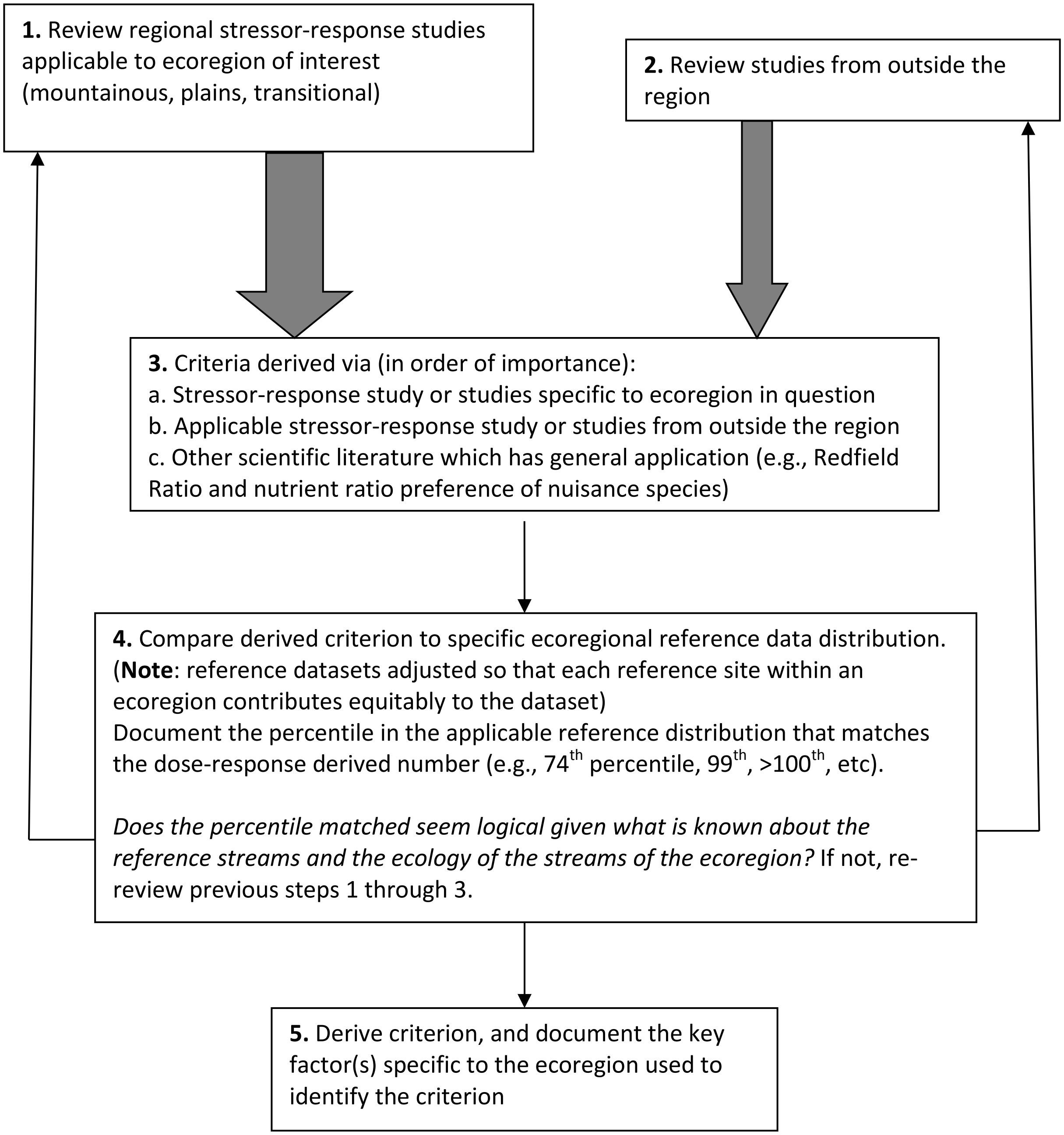

Figure 1-2. Overview of approach used to derive nutrient criteria in this document.
The size of the large grey arrows near the top of the figure represent the relative importance of the two information sources for deriving regional nutrient criteria.

The other major change, relative to the 2008 document, is that the Department will not be recommending nitrate (or nitrate + nitrite) criteria for adoption for the control of eutrophication at this time. Only total phosphorus (TP) and total nitrogen (TN) criteria are provided here. Rapid uptake of soluble nitrogen compounds by aquatic organisms (mainly algae and plants) makes these compounds' concentrations highly variable, and difficult to use as ambient surface water criteria. Total nitrogen and TP have been shown to provide better overall correlation to eutrophication response than soluble nutrients (Dodds et al., 1997; Dodds et al., 2006; Dodds et al., 2002) and, in terms of water quality criteria, total nutrients are more practical than soluble forms for river monitoring and assessment, total maximum daily loads, etc. (Dodd and Welch, 2000). **However**, the Department strongly encourages the collection of nitrate + nitrite when collecting TN and TP data. The soluble data can often point to specific types of nutrient sources, for example. The Department's Water Quality Monitoring Section will continue to include nitrate + nitrite alongside TN and TP for routine monitoring for nutrients and may use some general guidelines from the scientific literature for determining when measured concentrations are clearly too high.

2.0 METHODS USED TO DERIVE THE CRITERIA

In the **Introduction** we presented a general overview of the updated process used to derive the numeric nutrient criteria (**Figure 1-2**). In this section, we delve further into the details of these approaches.

2.1 ECOREGIONS AS THE BASIS FOR ESTABLISHING NUTRIENT CRITERIA ZONES

The Department tested the usefulness of ecoregions (Omernik, 1987) as a means to establish nutrient criteria zones; that work is detailed in Varghese and Cleland (2005) and Section 4.0 of Suplee et al. (2008). The Department will continue to use ecoregions as the basis for establishing nutrient criteria zones. Subsequent analysis has further verified that specific level IV (small scale) ecoregions are significantly different from the larger-scale level III ecoregions in which they reside (Varghese and Cleland, 2008; Varghese and Cleland, 2009). In **Section 3.0** of this document we will detail the criteria derived for individual ecoregions at the level III or level IV scale. In general, a level IV ecoregion will only be broken out for nutrient criteria derivation **if** (1) natural concentrations of nutrients in the level IV ecoregion are elevated above concentrations identified as harming uses per the stressor-response studies pertaining to that region, **or** (2) it is a level IV ecoregion that resides along the Rocky Mountain front (or similar environments) and represents a zone containing mountain-to-prairie transitional streams. In some cases the effect of a particular level IV ecoregion will influence natural nutrient concentrations in downstream waterbodies outside of the boundaries of the level IV. In **Section 4.0** a method to account for this type of influence is detailed and a number of reach-specific criteria are recommended there.

2.2 CRITERIA IN THIS DOCUMENT APPLY TO WADEABLE STREAMS

The scope of the criteria in this document is wadeable streams. The only substantive change since 2008 pertaining to this topic is the definition of specific rivers and river segments which are not wadeable (i.e., the large rivers). Flynn and Suplee (2010) use a wadeability index (product of river depth [in feet] and mean velocity [in ft/sec]) of 7.2 to segregate wadeable from non-wadeable rivers. During summer base flow these large rivers have mean water depths in excess of 3.15 ft and discharges of 1,500 ft^3/sec or greater. In Montana, rivers with these characteristics are almost always 7th order or higher (Strahler, 1964), and this is consistent with earlier definitions of large rivers based on stream order (Welcomme, 1985). **Table 2-1** shows the non-wadeable large rivers of the state to which the criteria in this document ***do not*** apply. The Department is primarily using process-based mechanistic water quality models to identify criteria for large river segments.

Table 2-1. Large river segments within the state of Montana.

River Name	Segment Description
Big Horn River	Yellowtail Dam to mouth
Clark Fork River	Bitterroot River to state-line
Flathead River	Origin to mouth
Kootenai River	Libby Dam to state-line
Madison River	Ennis Lake to mouth
Missouri River	Origin to state-line
South Fork Flathead River	Hungry Horse Dam to mouth
Yellowstone River	State-line to state-line

2.3 CRITERIA APPLY SEASONALLY, WITH EXCEPTIONS

As before, we recommend that the numeric nutrient criteria for wadeable streams and rivers apply seasonally, during that period when algae growth is peak and ensuing water quality impacts are maximal (i.e., the "Growing Season"). See **Table 2-2** below. For monitoring and assessment purposes, however, a ten day window (plus/minus) on the Growing Season start and end dates is acceptable, in order to accommodate year-specific conditions (e.g., an early-ending spring runoff). Best professional judgment is required to decide if early or later sampling is warranted.

Table 2-2. Start and Ending Dates for Three Seasons (Winter, Runoff and Growing), by Level III Ecoregion.

Ecoregion Name	Start of Winter	End of Winter	Start of Runoff	End of Runoff	Start of Growing Season	End of Growing Season
Canadian Rockies	Oct.1	April 14	April 15	June 30	July 1	Sept. 30
Northern Rockies	Oct.1	March 31	April 1	June 30	July 1	Sept. 30
Idaho Batholith	Oct.1	April 14	April 15	June 30	July 1	Sept. 30
Middle Rockies	Oct.1	April 14	April 15	June 30	July 1	Sept. 30
Northwestern Glaciated Plains	Oct.1	March 14	March 15	June 15	June 16	Sept. 30
Northwestern Great Plains	Oct.1	Feb. 29	March 1	June 30	July 1	Sept. 30
Wyoming Basin	Oct.1	April 14	April 15	June 30	July 1	Sept. 30

Exceptions to the seasonal applicability of nutrient standards will occur when it is known or demonstrated that a stream or river is having a significant influence on a downstream lentic waterbody (lake, reservoir). In such cases, criteria (and nutrient loads) applicable to the lake may apply to a stream draining to the lake, and would apply year round. These situations need to be determined case-by-case, and are beyond the scope of this document.

2.4 METHOD FOR ASSURING THAT ALL REFERENCE SITES ARE EQUITABLY REPRESENTED IN AN ECOREGION (THE ALL-OBSERVATIONS DATASET AFTER APPLYING BRILLOUIN EVENNESS INDEX, AND THE MEDIAN DATASET)

Assuring that each reference site contributes an approximately equal number of N and P observations to each ecoregional zone has been a Department objective for some years (see Section 6.2.1 of Suplee, et al. (2008)). Since Suplee et al. (2008) was released, the Department continued to target under-represented reference sites and collected data in summer 2009 and summer 2010.

In spite of the targeted field work, there was still a fair amount of inequality in terms of the number of nutrient observations per site in each ecoregion. Therefore, we undertook two different methods in the office using the updated (current through 2010) reference nutrient dataset. In method one we used (as in 2008) the Brillouin evenness index (Pielou, 1966; Zarr, 1999). This method assures that each reference site in an ecoregion contributes equal amounts of information to the nutrient dataset. Distribution statistics (e.g., maximum, median, 75th percentile) are calculated from the resulting dataset, and these statistics provide a means for readers to see the full range of nutrient concentrations that have been observed across reference sites during the growing season. The method assumes each observation from a reference site is independent of the others. Nutrient samples are collected from reference sites in a way intended to maximize independence, and a number of tested case studies show independence is usually maintained (Suplee and Sada de Suplee, 2011). In this document datasets so handled are called "all-observations after applying Brillouin Evenness Index", or simply the "all observations dataset".

In method two, all the growing-season observations from a reference site within an ecoregion are first reduced to a median. Distribution statistics are then calculated on the population of site medians; these are the "median datasets" in this document. Method two addresses the potential for intra-site temporal pseudoreplication (Hurlbert, 1984), but the output masks the full range of nutrient concentrations that have been observed across the ecoregion's reference sites. For this reason only the interquartile range and the 90th percentile are reported for the median datasets.

Because method one is computationally involved, it is detailed here. The Brillouin evenness index (J) for a whole population (Pielou, 1966; Zarr, 1999) is:

$$\mathbf{J} = H \div H_{max}$$

with

$$H = \frac{(\log n! - \sum \log f_i!)}{n}$$

and

$$H_{max} = \frac{\log n! - (k - d) \log c! - d \log (c + 1)}{n}$$

where n is the total number of reference nutrient observations (e.g., TP) in the ecoregion, f_i is the frequency of nutrient observations specific to each reference site in the ecoregion, k is the number of reference sites in the ecoregion, c is the integer portion of n/k and d is the remainder. The index value J will range from zero to one (one being the case where each reference site has been sampled for nutrients exactly the same number of times).

We wanted to achieve an evenness of 90% or better for each ecoregional reference dataset. Applying these equations with a target J value of ≥ 0.9 required in some cases that a proportion of observations from heavily-sampled reference sites be excluded from use. This was carried out objectively and independently for TN and for TP, as follows. First, the J value was calculated using all data for a given nutrient (e.g., TP) from all reference sites within the ecoregion in question. If the value was ≥ 0.9, nothing further was done and all the data were used as-is for descriptive statistics. If the value was <0.9, we identified the over-contributing reference sites and calculated how many observations would need to be removed from each in order to achieve a J value of 0.9. Because J measures evenness, reducing many observations from a single over-contributing site was not effective. Instead, a smaller number of observations had to be eliminated from each of the major over-contributors. Once the number of observations to be eliminated from each over-contributing reference site was known, we randomly removed that number of observations from the dataset of each of the specified sites. Finally, the now 'more even' ecoregional dataset (comprising the remaining observations from sites where the random-elimination process was applied plus the observations from the sites where no censoring was applied) was used to generate descriptive statistics.

In **Sections 3.0** and **4.0** we present distribution statistics (e.g., 25th, 75th percentiles) for data processed by both methods one and two. The results are sometimes different for each, and this is a function of the way the data were processed. It is important to note that no inferential statistics were carried out nor are the recommended criteria tied to any specific percentile; the distributional qualities of the reference data are being provided primarily as a means for readers to compare the recommended criteria to regional reference data. It should also be noted that in some level IV (small scale) ecoregions, there were only a few references sites and the number of collected nutrient samples was correspondingly low. We wanted to maintain a sample-size minimum of about 12 (Varghese and Cleland, 2008, Appendix H) for these ecoregions in order to sufficiently characterize the reference condition. Therefore, in level IVs that were near to this sample-size minimum, no sample-size reductions using the Brillouin evenness index were undertaken.

2.5 LITERATURE CONSULTED

A re-review of the relevant scientific literature cited in Suplee et al. (2008) was undertaken, as well as a search and review of various studies and reports that have been released before and since 2008. The Department completed a whole-stream nitrogen and phosphorus addition study between 2009 and 2011 (Montana Department Environmental Quality, 2009). Findings from that study are incorporated into this work as well. The Department also completed a mechanistic water quality model (QUAL2K) for the lower Yellowstone River and has recommended criteria for that waterbody using the model (Flynn and Suplee, 2013). Although the later work pertains to large rivers, findings from it help define the range of nutrient criteria one might expect for flowing waters of Montana.

Details on the specific literature that was most useful within each ecoregion will be provided in **Section 3.0.**

2.5.1 Literature Pertaining to Nutrient Enhancements in Rivers and Streams

Much of the pertinent scientific literature of the past few decades focuses on the effects of nitrogen (N) and phosphorus (P) over-enrichment. However, there is a smaller but equally valuable body of scientific literature addressing intentional nutrient *additions* to rivers and streams; these actions have usually been carried out for the purpose of enhancing depleted fisheries production (Holderman et al., 2009; Stockner, 2003). As we pointed out in Section 1.3 of Suplee et. al (2008), N and P have an interesting duality in that too much is a problem (cultural eutrophication), but too little can also be a problem (cultural oligotrophication). Thus, the nutrient-addition literature enabled us to have a better understanding of the ecology of nutrient-poor rivers and streams and how that ecology shifts as nutrients increase towards the concentrations that ultimately become "too much of a good thing". Many of the nutrient-addition studies were carried out in the Pacific Northwest in streams and small rivers similar to those found in western Montana (Perrin et al., 1987; Johnston et al., 1990; KOHLER et al., 2008; Perrin and Richardson, 1997; Stockner and Shortreed, 1978). Because salmon die after spawning in the upper tributaries of rivers draining to the Pacific, large quantities of marine-sourced nutrients are relocated to the streams annually. But overfishing, dams, and habitat destruction have greatly reduced many salmon runs, leaving the streams stripped of their annual nutrient source. To boost survival of the few fry and fingerlings that are spawned, nutrient additions have been undertaken by resource managers. These nutrient additions enhance algal growth and secondary production (aquatic insects), which in turn provide a larger food source for the fish which enhances their growth and survival (Stockner and Ashley, 2003). This type of work has included large-scale nutrient additions to the Kootenai River as it flows out of Montana into Idaho. Ambient nutrients in the Kootenai River were greatly reduced after the completion of Libby dam in the early 1970s (Holderman et al., 2009).

The streams to which nutrients are added for fisheries enhancement have very low ambient nutrient levels (ca. 5-10 µg TP/L and < 15 µg NO_3-N/L), and nutrient concentrations are only increased by a few additional micrograms per liter. The studies reviewed provided good incite on the ecological changes that occur in low-nutrient streams once nutrients increased, and how salmonid fisheries react to these small nutrient increases.

2.6 BOTH NITROGEN AND PHOSPHORUS CRITERIA ARE RECOMMENDED

The concept of nutrient limitation is important in the development of N and P criteria. Relative to N and P, limitation can be defined in a negative sense; a nutrient is *not* limiting if, when increased, one does not observe an effect on plant or algal growth (Gibson, 1971). The scientific literature has many examples of studies and analyses showing that N, or P, or commonly both stimulate algal production in surface waters (Francoeur, 2001; Smith et al., 1999; Tank and Dodds, 2003; Elser et al., 1990; Elser et al., 2007; Lewis et al., 2011). Co-limitation appears to be especially common in flowing waters , where nutrient-addition experiments show that added N and P result in much greater response of algal growth than does N- or P-addition alone (Elser et al., 2007). Regional work using nutrient diffusing substrates (N, P, and N+P) supports these findings (Mebane et al., 2009). Mebane et al.'s experiments were carried out *in situ* in intermontane wadeable streams of Idaho which are comparable to Montana's western streams. Background N and P concentrations in the streams ranged from very low (7 µg TP/L and 50 µg TN/L) to quite elevated (e.g., 91 µg TP/L and 1,820 µg TN/L). Based on the growth of algae on the nutrient diffusers that developed over 21 days, N and P co-limitation was indicated in three streams, N limitation was shown in two streams, and P limitation was found in one stream; one stream with highly elevated ambient nutrients showed no limitation (Mebane et al., 2009). And it should be noted, especially in light of the definition of nutrient limitation given above, that in most of the streams the greatest algal biomass developed on the N+P diffusers (Mebane et al., 2009).

Liebig's Law of the Minimum (Hooker, 1917) is a well-established tenet in the agricultural sciences that states that biomass yield for a particular plant is usually limited by the nutrient that is present in the environment in the least quantity relative to the plant's need for that nutrient to support growth. The law is sometimes used to rationalize the idea that, in most cases, only P needs to be reduced to low concentrations to achieve eutrophication control in freshwaters. But Liebig's Law best applies to single plant species at a given place at a certain time, whereas the numeric nutrient criteria in this document apply to a mixed flora in flowing waters over several months of growing season. These flowing waters receive variable N and P loads over time and are home to mixed populations of algae species—and each species has somewhat different N and P requirements and capability of taking up nutrients (Hecky and Kilham, 1988; Borchardt, 1996). Nutrient limitation of the aggregate algal community is largely a function of the nutrient limitation of the dominant species, but shifting nutrient availability can change the dominant species and potentially the limiting nutrient.

Streams are variable environments where, for example, N and P availability can alternate as a function of stream discharge (Hullar and Vestal, 1989). Stated simply, limiting nutrient levels are not fixed and both nutrients are likely to limit some facet of the algal community at any point in time. If for example P is presently limiting in a stream, that does not mean there is no point in limiting N. If P were to increase, say from summer rain events, or due to the confluence of a downstream tributary with slightly higher P concentrations, the N that was formerly in excess can become the limiting nutrient without any change in its absolute concentration (Gibson, 1971). Similarly, an algal community may be N-limited early in summer, and as surface flows drop, proportionally more N-rich groundwater enters the stream, shifting the community structure and switching the stream to P-limitation. Results from twelve years of monitoring on the Clark Fork River in Montana support the idea that it is best to control both N and P. There, in river locations where both the N standard and the P standard were met (20 µg TP/L and 300 µg

TN/L); algal biomass has usually been reduced to the standard (150 mg Chl*a*/m^2). Locations in the river where these nutrient levels have not been met continue to have elevated algae biomass, and study sites give mixed signals regarding nutrient limitation—some suggesting N limitation, others P; these signals are not consistent across time or location (Suplee et al., 2012).

Water quality standards based on control of only a single nutrient (i.e., P) could result in unwanted ecological consequences in Montana's rivers and streams. Background nutrient levels in our western reference streams are usually quite low (10-18 µg TP/L and 85-190 µg TN/L; Smith et al., 2003; Suplee et al., 2007), and usually have TN:TP ratios at or somewhat higher than Redfield (Redfield = 7:1 by mass; Redfield, 1958). The nuisance diatom alga *Didymosphenia geminata* has, in recent years, spread to and formed nuisance benthic blooms in low-nutrient rivers and streams worldwide (Whitton et al., 2009; Kilroy, 2011; Spaulding and Elwell, 2007). It is found in Montana and, in western U.S. states, probabilistic survey data show that in over half of streams containing *D. geminata* TP is <10 µg/L (Spaulding and Elwell, 2007). Others also report that *D. geminata* usually occurs in streams with very low P (Whitton et al., 2009; Kilroy and Bothwell, 2012), and that it tends to disappear when TP exceeds about 20 µg/L (Lovstad, 2008). Further, *D. geminata* generally thrives in waters where N:P ratios are high (34:1 on average) much of the time (Whitton et al., 2009). *Didymosphenia geminata* blooms in low-P streams are caused by the diatoms' elevated production of polysaccharide stalks which develop as a consequence of phosphorus limitation (Kilroy, 2011; Kilroy and Bothwell, 2012). Stalk production in attached diatoms is considered competitively advantageous because it elevates the cells towards higher light (Hudon and Bourget, 1981), and this also places them in closer contact with available nutrients in flowing water (Kilroy and Bothwell, 2012; Bothwell et al., 2012)[3].

Researchers suggest that an effective way to diminish *D. geminata* blooms is to encourage its algal competitors by assuring that a sufficient (though small) supply of soluble phosphorus is available (Whitton et al., 2009; Kilroy and Bothwell, 2012). Indeed, the Montana Department of Fish, Wildlife and Parks is currently planning low-level phosphate addition experiments in troughs alongside the Kootenai River (where *D. geminata* blooms have become quite severe) to see if the alga can be brought under control via nutrient management actions. The alga is believed to be impacting the salmonid fishery there, where the high algal density appears to be reducing abundance of key aquatic insects which salmonids prey upon (Jim Dunnigan, Fishery Biologist, MT Fish Wildlife and Park, personal communication March 14, 2012).

---

[3] Alternative hypotheses exist regarding what controls *D. geminata* blooms in low-P streams. One of them states that the diatoms' polysaccharide stalks have an affinity for iron, which is absorbed to the stalks and in the process forms iron oxyhydroxide (Sundareshwar et al., 2011). Iron oxyhydroxides have a strong affinity for P and will adsorb (co-precipitate) it from the water (Mortimer, 1941; Caraco et al., 1989; Hasler and Einsele, 1948). Sundareshwar et al. (2011) posit that as the mat grows, anaerobic microbial decomposition (by iron- or sulfate-reducing bacteria) of dead diatoms within the mat leads to the reduction of the iron, formation of iron sulfides, and concomitant release of P. The abundant P is then available to the live diatoms at the mat surface, supporting further growth. The geochemical process they describe is well known in marine and freshwater systems (Suplee and Cotner, 2002). But the mechanism that Sundareshwar et al. (2011) propose is unsatisfactory, as it does not explain why the mats grow rapidly and develop to great size in low P streams in the first place. Sundareshwar et al. (2011) note that *D. geminata* produces high levels of alkaline phosphatase at the mat surface because the P sequestered there with iron is *not* bioavailable. Thus, we view their hypothesis as a potential mechanism for mat maintenance, but not necessarily for mat development. Others also find that this geochemical explanation does not jive with findings in *D. geminata* dominated streams, where increases in stream P concentrations lead to *declines* in *D. geminata* (Bothwell et al., 2012; Kilroy and Bothwell, 2012).

Phosphorus reduction is often promoted as the only eutrophication control approach needed to control eutrophication, this being based largely on the more economical removal of P from wastewater (Lewis and Wurtsbaugh, 2008), and the assumption that P can be made to become limiting in the waterbody, *sensu* Liebig's Law. But the facts given above indicate that it would be unwise to recommend only P standards for control of excess algal biomass in our streams and rivers. A P-only approach, in order to work, would require that P standards be set to the background levels observed in our western region's reference sites (e.g., 10 µg TP/L). Total phosphorus concentrations this low are hard to achieve technologically, but if the P standard was not set to this low natural background, then the commonly occurring N-limitation or N and P co-limitation would lead to algal growth stimulation nonetheless. Worse yet, in the long term, a P-only strategy would result in highly skewed (elevated) N:P ratios accompanying the low P levels. These management-induced conditions might control green algae biomass but may lead to nuisance blooms of *D. geminata.*

A balanced and prudent policy would be to reduce both N and P and maintain, as nutrient concentration reductions occur, a roughly balanced (i.e., Redfield) ratio between the two. This is the strategy that has been applied on the Clark Fork River and it appears to be working (Suplee et al., 2012). Other researchers in the field have recommended that both N and P need to be controlled to effectively manage eutrophication (Conley et al., 2009; Lewis et al., 2011; Paerl, 2009). Thus, we will generally be recommending both N and P criteria for wadeable streams and rivers in this document.

One final word on Redfield ratios. Studies of benthic algae show that it is necessary to move some distance above or below the Redfield ratio in order to be strongly convinced that a lotic waterbody is P or N limited (Dodds, 2003). When a benthic algal Redfield ratio (by mass) is <6, N limitation is suggested, and when it is >10 P limitation is indicated (Hillebrand and Sommer, 1999). Thus, there is a range of N:P values between about 6 and 10 where one can state, for practical purposes, that algal growth is co-limited by N and P.

3.0 ECOREGION-SPECIFIC NUMERIC NUTRIENT CRITERIA RECOMMENDATIONS

In this section are documented the numeric nutrient criteria for each ecoregion. Ecoregional information is arranged as follows: (1) first the level III ecoregion is presented, and (2) if any level IV ecoregions within the level III need to be treated separately, their information follows in a subsection. The same presentation format is followed for each ecoregion, be it level III or level IV, to the degree possible. Data and discussion specific to each ecoregion is presented on three or four pages. A map of Montana showing the ecoregion in which the criteria apply is shown first, followed by criteria recommendations and tables of descriptive statistics for the reference sites in the ecoregion. **The Redfield ratio shown for the reference sites is based on the 50th percentile from the applicable median dataset.** Then readers will find: histograms of the reference data TN and TP distributions based on the all-observations dataset evened using the Brillouin index (in cases where the data were skewed to the right they have been $\log_{10}$ transformed); a discussion of the scientific studies (regional and beyond) that were used to help derive the criteria, any other considerations pertaining to the derivation of the criteria; and a conclusion summarizing final thoughts on the criteria.

Data from reference sites (Suplee et al., 2005) were important in the process of deriving the nutrient criteria (see **Figure 1-2**). **Figure 3-1** below is a statewide map showing the locations of all stream reference sites current through August 2011. There are currently 185 different sites in the network.

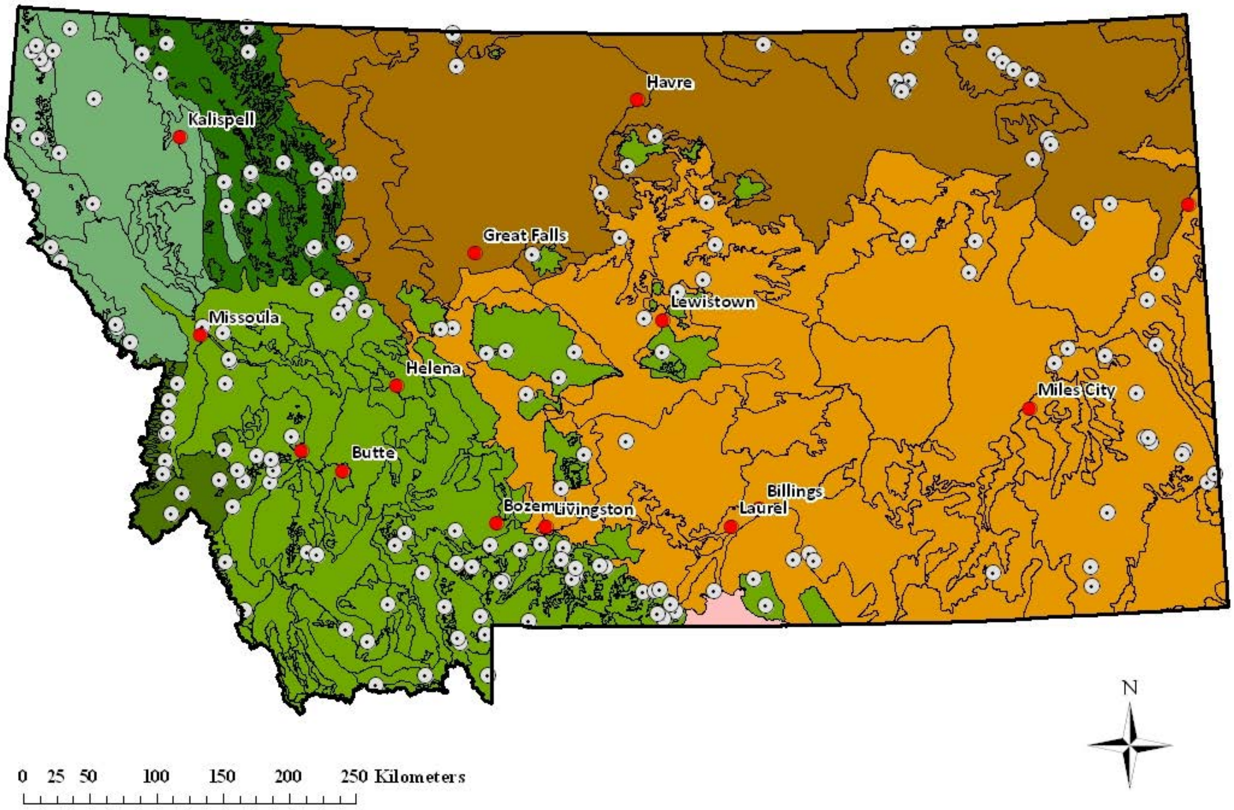

Figure 3-1. Map of Montana showing location of stream references sites (white dots). Colored regions denote level III ecoregions. Red dots show the major towns.

3.1 Level III: Middle Rockies (Ecoregion 17)

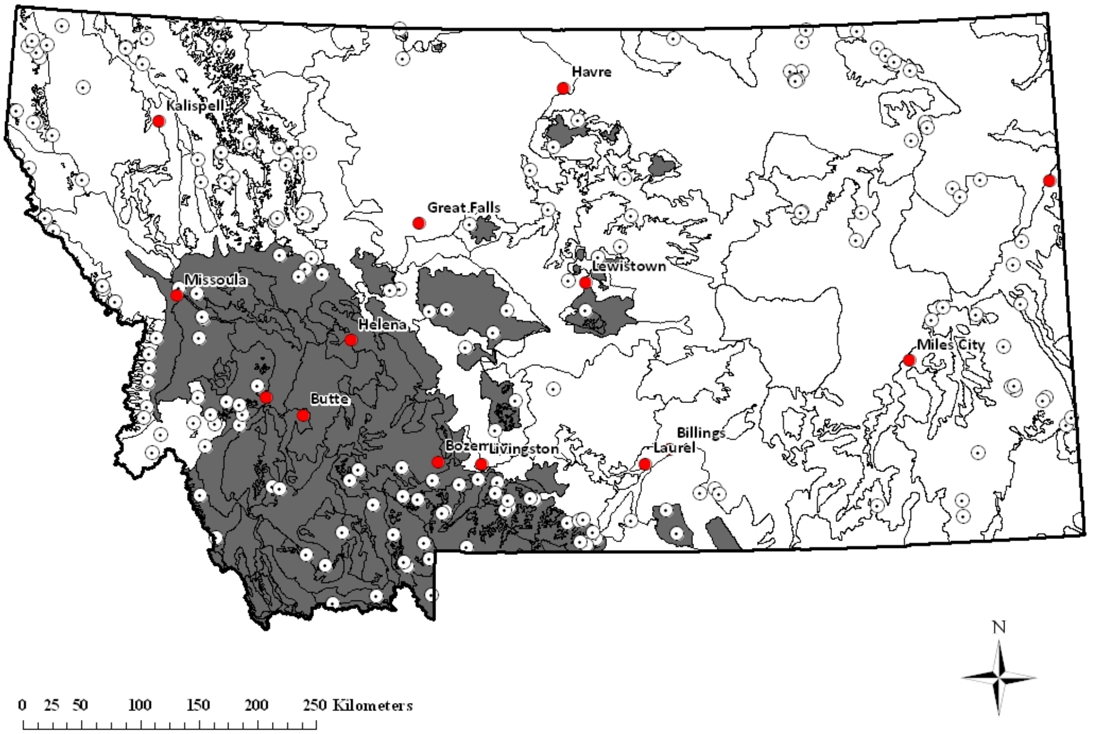

Figure 3-2. Map of Montana showing the Middle Rockies ecoregion in gray.
White dots are the reference sites.

Recommended Numeric Criteria

Total Phosphorus: **30 µg TP/L**
Total Nitrogen: **300 µg TN/L**

N:P Ratio of Criteria: **10:1**
N:P Ratio of Reference Sites: **11:1** (Redfield N:P ratio = 7:1)

Descriptive Statistics of Regional Reference Sites

Table 3-1A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Middle Rockies ecoregion.
Data are from the all-observations dataset after applying Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	57	148	3	9580	55	95	141	220
TP	61	245	0.5	840	6	10	20	70

Table 3-1B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Middle Rockies ecoregion.
Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	57	51	90	136	181
TP	61	4	8	15	43

Criteria Match to Reference Distributions:

The 30 µg TP/L criterion matches to the 80th percentile of the all observations dataset and the 82nd percentile in the median dataset.

The 300 µg TN/L criterion matches to the 93th percentile of reference of the all observations dataset and the 98th percentile in the median dataset.

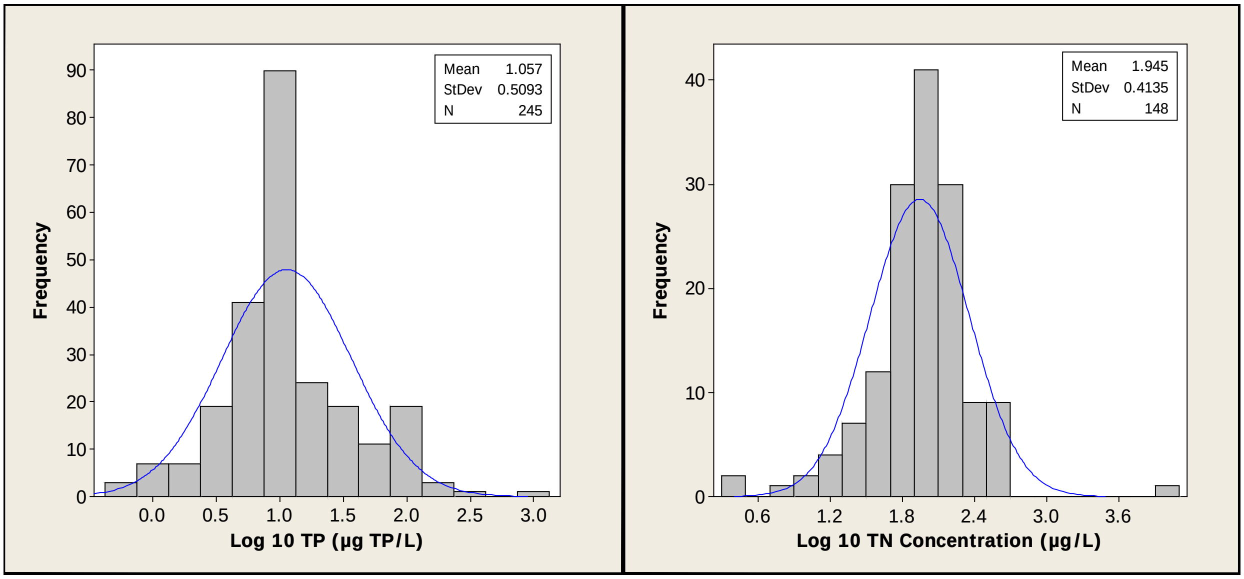

Figure 3-3. Nutrient concentrations from reference streams in the Middle Rockies ecoregion.
Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). Data were collected during the Growing Season (July 1-September 30).

Discussion of the Middle Rockies Ecoregion Nutrient Criteria

Two regional dose-response studies were available that relate TN and TP to stream impacts (Mebane et al., 2009; Suplee et al., 2012). Suplee et al. (2012) show that TP is saturated in the Clark Fork River at 24 µg/L. They also suggest that criteria for the Clark Fork River— upstream of the Flathead River confluence— be set uniformly to about 20 µg TP/L and 300 µg TN/L to meet the algae standard (150 mg Chl*a*/m^2 max). Further, they indicate that both TN and TP criteria should be met to achieve the intended reductions in algal biomass. Suplee et al. (2012) build on earlier work in which nutrient criteria were developed for the Clark Fork River (Dodds et al., 1997), and by doing so provide large-scale confirmation that the original criteria were largely correct. Mebane et al. (2009) carried out a study in southern Idaho. Many of the streams were intermontane and, thus, very similar to intermontane streams of this ecoregion. They recommend 40 µg TP/L and 600 µg TN/L in order to maintain benthic algae growth ≤150 mg Chl*a*/m^2, per Suplee et al. (2009). To maintain 125 mg Chl*a*/m^2, which is generally appropriate for shallow wadeable streams[4], the values would drop to about 35 µg TP/L and 475 µg TN/L.

Beyond the Middle Rockies ecoregion, studies in northern and southern temperate rivers and streams show that nutrient-benthic Chl*a* regressions have breakpoints at 27-62 µg/L for TP and between 367-602 µg/L for TN (Dodds et al., 2006; Dodds et al., 2002). What this indicates is that above the breakpoint concentrations, nutrients are saturated, and benthic algae control via nutrient control become ineffective. Stevenson et al. (2006) show in Michigan streams that the likelihood of reaching bottom coverage by *Cladophor*a of 20-40% increases sharply when TP exceed 30 µg/L and TN exceeds 1,000 µg/L. This level of streambed coverage by *Cladophora* was found to be unacceptable to the Montana public (see Suplee et al., 2009, Table 1). Chambers et al. (2011) derive nutrient criteria for Canadian streams using multiple methods including dose-response relationships between nutrients and algae and macroinvertebrate metrics. They recommend 20 µg TP/L and 210 µg TN/L for the Montane Cordillera, a mountainous region in British Columbia (actually, part of the Northern Rockies ecoregion). Equations relating benthic algal Chl*a* to total nutrients (Dodds et al., 1997, Equation 17; Dodds et al., 2006; Equation 19), were used to calculate TN levels that would maintain 125 mg Chl*a*/m^2 benthic algae given a TP of 30 µg TP/L. These equations resulted in TN concentrations ranging from 466-718 µg TN/L. If the algae level is set instead to 150 mg Chl*a*/m^2, and 30 µg TP/L is again used, the TN values range from 750 to 1,210 µg/L.

Conclusion

Studies that have the most specificity to the Middle Rockies suggest criteria ranging from 20-40 µg TP/L and 300-600 µg TN/L. Studies further afield provide a range of criteria to prevent nuisance algal growth or impacts to aquatic life communities ranging from 20-30 µg TP/L and 210-1,210 µg TN/L. **We recommend for this ecoregion 30 µg TP/L and 300 µg TN/L**. We recommend these values because: (1) these concentrations fall within the ranges provided in the studies, especially studies that are most pertinent to the ecoregion; (2) they maintain an N:P ratio of 10 which is close to the natural condition of regional reference sites (i.e., 11:1) and is fairly close to the upper band of the Redfield ratio and that indicates slight P limitation; and (3) they should generally encourage a balanced and diverse stream flora for this region by keeping nutrient ratios not far from Redfield and TP at a concentration which will help inhibit *Didymosphenia geminata* blooms.

---

[4] A nutrient dose-response study carried out by the Department in southeastern Montana showed that in a wadeable stream benthic algae levels of 127 mg Chl*a*/m^2 (33 g AFDM/m^2) led to seasonal exceedances of the dissolved oxygen (DO) standard (Suplee and Sada de Suplee, 2011). Subsequent work—using a model based on Streeter-Phelps (1925) and cooler water temperatures more typical of mountainous streams—showed that DO exceedances would still occur in many (though not all) western MT streams. Therefore, when using Chl*a*-nutrient relationships from Mebane et al. (2009), Dodds et al. (1997; 2006), and others, we also used 125 mg Chl*a*/m^2 as the target algae level. The Department believes this value is a well-supported threshold, giving consideration to both the DO impacts observed in the dosing study and also the recreational threshold (and regarding the later, giving consideration to the known statistical patterns provided in the Department's SOP Chl*a* method). In this document we continue to use 150 mg Chl*a*/m^2 as well, which corresponds to the arithmetic mean of the replicates at the highest level of benthic algae found to be acceptable to the MT public (Suplee et al., 2009).

3.1.1 Level IV Ecoregion within the Middle Rockies: Absaroka-Gallatin Volcanic Mountains (17i)

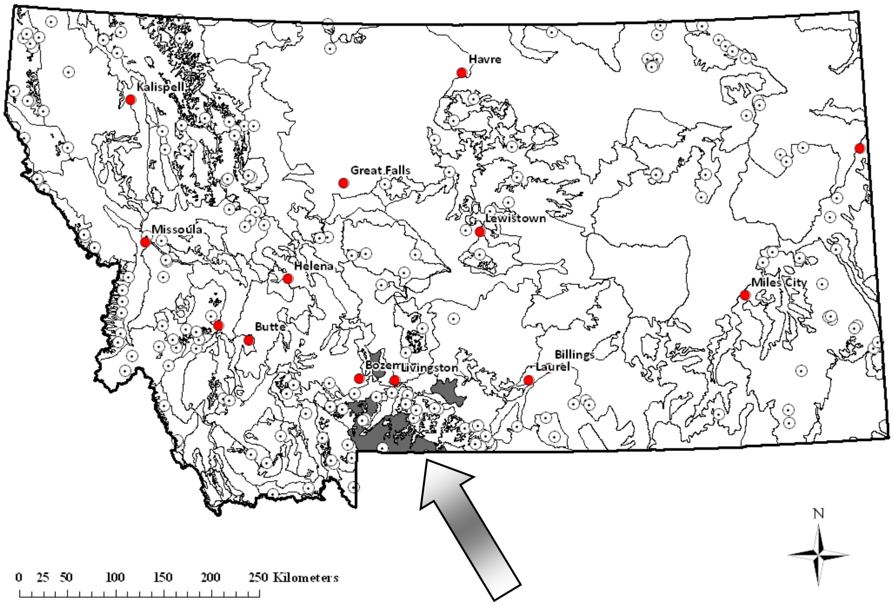

Figure 3-4. Map of Montana showing the Absaroka-Gallatin volcanic Mountains (17i), a level IV ecoregion within the Middle Rockies ecoregion.
White dots are the reference sites.

Recommended Numeric Criteria
Total Phosphorus: **105 µg TP/L**
Total Nitrogen: **250 µg TN/L**

N:P Ratio of Criteria: **2:1**
N:P Ratio of Reference Sites: **0.8:1** (Redfield N:P ratio = 7:1)

Descriptive Statistics of Regional Reference Sites

Table 3-2A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Absaroka-Gallatin Volcanic Mountains (17i) ecoregion.
Data are from the all-observations dataset after applying the Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	4	13	7	181	52	80	100	163
TP	4	16	16	144	61	81	105	127

Table 3-2B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Absaroka-Gallatin Volcanic Mountains level-IV ecoregion.
Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	4	42	65	83	93
TP	4	62	77	90	106

The 105 µg TP/L criterion matches to the 75th percentile of reference in the all observations dataset and the 89th percentile of reference in the median dataset.

The 250 µg TN/L criterion is greater than the 100th percentile of reference in both datasets.

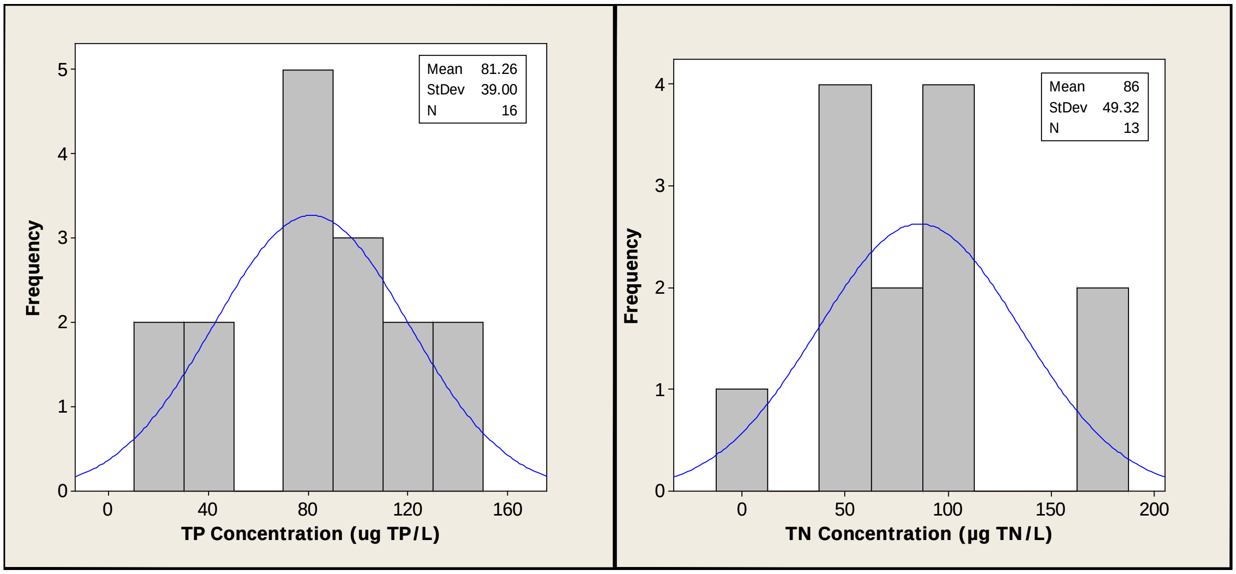

Figure 3-5. Nutrient concentrations from reference streams in the Absaroka-Gallatin Volcanic Mountains (17i) ecoregion.

Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). Data are from the Growing Season (July 1-September 30).

Discussion of the Absaroka-Gallatin Volcanic Mountains Ecoregion

The Absaroka-Gallatin Volcanic Mountains ecoregion (17i) has statistically significantly higher TP concentrations than the rest of the Middle Rockies (Varghese and Cleland, 2008; Varghese and Cleland, 2009). Permian age Phosphoria formations (United States Geological Survey, 1951) outcrop throughout this ecoregion and cause naturally elevated P concentrations. The natural concentrations of TP in 17i exceed harm-to-use thresholds identified for the Middle Rockies (20-40 µg TP/L). The median TP concentration of reference streams in ecoregion 17i is 77 to 81 µg/L (median or all observations datasets, respectively), compared to 8-10 µg/L for the Middle Rockies as a whole, and is therefore already higher than saturation (Dodds et al., 2006).

Observation of the reference streams of this ecoregion indicate that nuisance levels of benthic algae are not developing. This suggests that they are N limited, otherwise one would expect high algae levels at these TP concentrations (as observed in the transitional level IV ecoregions of the Northwestern Glaciated Plains, discussed later on). Natural TN levels in these streams are fairly low, lower than what is observed in the Middle Rockies as a whole. To assure that management-induced changes in TN do not lead to stream impacts, careful consideration of the appropriate TN criterion was essential. Equations relating benthic algal Chl*a* to total nutrients (Dodds et al., 1997, Equation 17; Dodds et al., 2006; Equation 19) were used to calculate TN that would maintain 125 mg Chl*a*/m^2 benthic algae given a TP of 105 µg/L (105 µg TP/L = 75th percentile of reference of 17i). This resulted in TN concentrations from 245 to 287 µg TN/L. If the algae level is set instead to 150 mg Chl*a*/m^2, and 105 µg TP/L is again used, the TN values range from 322 to 483 µg/L. Total phosphorus at the 75th percentile of reference was selected because it assures that the majority of data from the ecoregion's reference sites are below the TP criteria, and it lends itself well to reach-specific criteria derivation in cases where a stream reach further down gradient receives water from both the Middle Rockies and the Absaroka-Gallatin Volcanic Mountains (more on this in **Section 4.2**).

Conclusion

We recommend 105 µg TP/L and 250 µg TN/L as criteria for this level IV ecoregion. The TN criterion is more restrictive here than the 300 µg/L recommended for the Middle Rockies, and more restrictive than the other western ecoregions that will be discussed below; this is to assure adequate control of N in these apparently N-limited streams. The criteria have an N:P ratio of 2:1, however this is acceptable because maintaining a ratio near to Redfield is not realistic (or necessary) since the streams' natural N:P ratios are already low (on the order of 1:1). The reference data for this ecoregion were collected between 1990 and 2009, providing good temporal dispersion. Since there are still only a minimal number of samples (13-16) available for characterizing this ecoregion we recommend continued sample collection to increase the sample size.

3.2 LEVEL III: NORTHERN ROCKIES (ECOREGION 15)

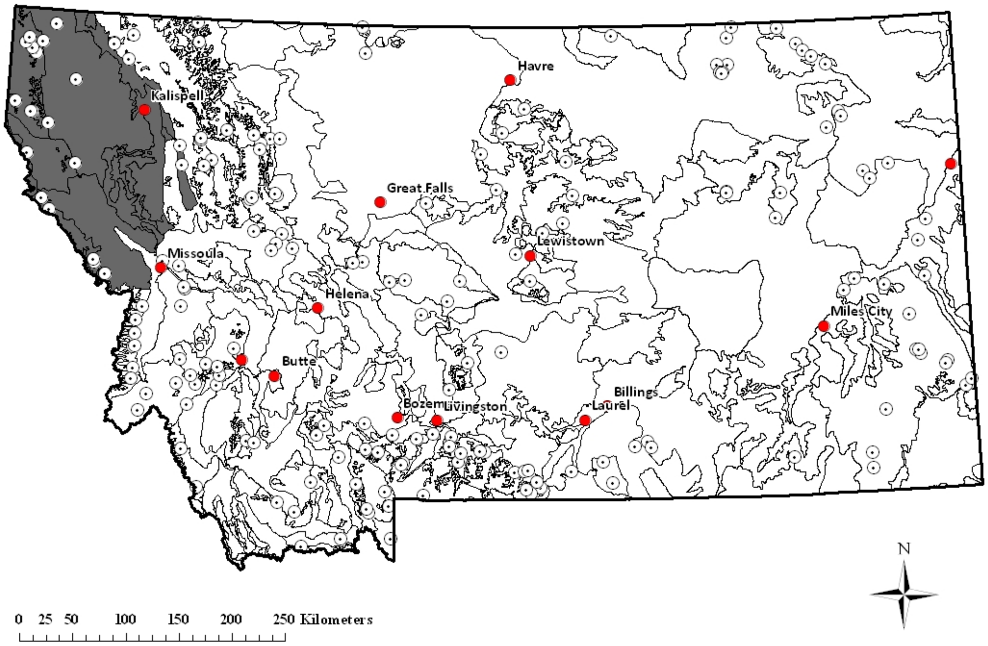

Figure 3-6. Map of Montana showing Northern Rockies ecoregion.
White dots are the reference sites.

Recommended Numeric Criteria

Total Phosphorus: **25 µg TP/L**
Total Nitrogen: **275 µg TN/L**

N:P Ratio of Criteria: **11:1**
N:P Ratio of Reference Sites: **8:1** (Redfield N:P ratio = 7:1)

Descriptive Statistics of Regional Reference Sites

Table 3-3A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northern Rockies ecoregion.
Data are from the all-observations dataset after applying Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	22	76	3	360	18	41	94	167
TP	22	81	0.5	18	4	6	9	13

Table 3-3B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northern Rockies ecoregion.
Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	22	18	39	79	131
TP	22	3	5	9	13

The 25 µg TP/L criterion is greater than the 100th percentile of reference for both datasets.
The 275 µg TN/L criterion matches to the 96th percentile of reference (all observations dataset) and the 99.5th percentile in the median dataset.

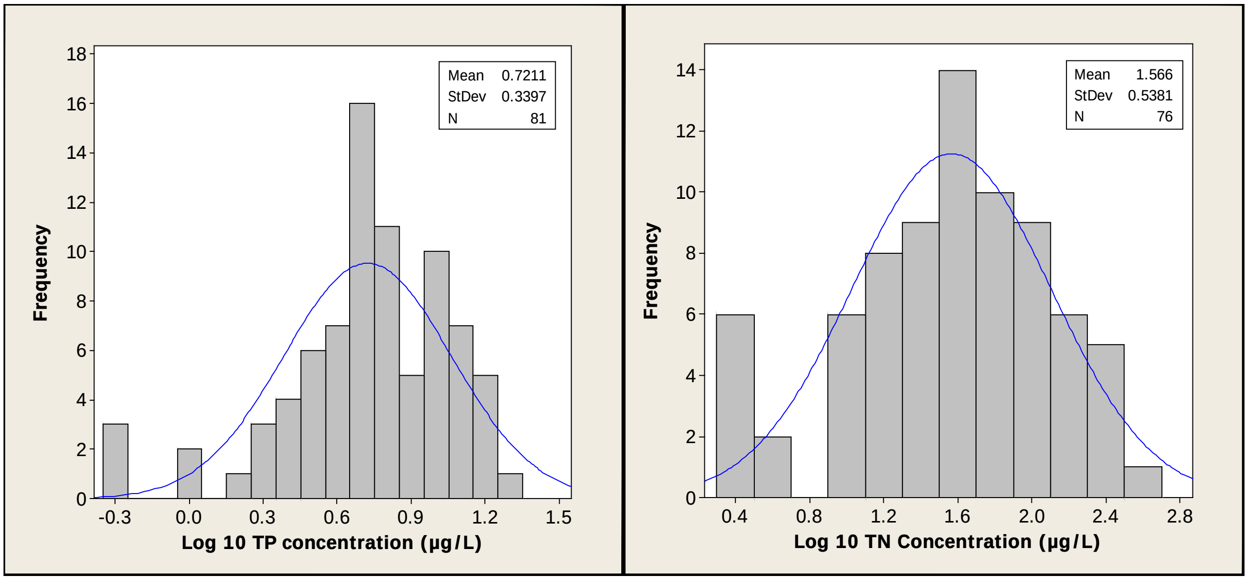

Figure 3-7. Nutrient concentrations from reference streams in the Northern Rockies ecoregion. Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). Data were collected during the Growing Season (July 1-September 30).

Discussion of the Northern Rockies Ecoregion Nutrient Criteria

Three regional dose-response studies specific to the Northern Rockies were available that relate nutrients (both soluble and total forms) to stream impacts or changes in aquatic communities. Welch et al. (1989) use a model and an artificial stream study and then adapt them to an open river system (Spokane River, Washington). Their equations indicate that at 10 µg soluble reactive phosphate (SRP)/L, the distance on the river with algal biomass of 150 mg Chl*a*/m^2 would be constrained to 16 km. The Montana public found a mean of ≤150 mg Chl*a*/m^2 acceptable for river recreation[5]. Assuming an SRP:TP ratio of 0.25:1 (as is commonly observed on the Clark Fork River), 10 µg SRP/L equals 40 µg TP/L. Chambers et al. (2011) derive nutrient criteria for Canadian streams using multiple methods including dose-response relationships between nutrients and algae and macroinvertebrate metrics. The study streams were located in the Okanagan Basin (British Columbia) just north of Washington State, and are within the Northern Rockies ecoregion. They recommend 20 µg TP/L and 210 µg TN/L for streams of the region to protect aquatic life.

The third study (Gravelle et al., 2009a; Gravelle et al., 2009b) discusses a Before After Control Impact Paired study in which the authors assess the effects of different timber harvest intensities on nutrient concentrations and aquatic insect metrics in the Mica Creek Experimental Watershed in northern Idaho.

---

[5] The Spokane River is a 6th order river and is therefore on the large side of wadeable (Flynn and Suplee, 2010), and impacts to dissolved oxygen standards would be less likely in a river this size due to good re-aeration and total river volume. Therefore we only used 150 mg Chl*a*/m^2 in the equation (as opposed to 125 mg Chl*a*/m^2, discussed in **footnote 4**).

In the post-road construction period (1998-2001), summer TP increased to about 40 µg/L, TKN increased slightly to about 150 µg/L, and nitrate+nitrite did not change. Later, in the post-harvest period (2002-2006), TP declined again to 20 µg/L and TKN to about 40 µg/L, but nitrate+nitrite increased markedly to a monthly summer average of 350 µg/L (about 400 µg TN/L). Across this entire ten year period there were very few changes in the aquatic insect metrics monitored, although Ephemeroptera, Plecoptera, and Trichoptera abundance increased over the period (Gravelle et al., 2009b). Among the biometrics, the Hilsenhoff Biotic Metric (HBI) was of particular interest as the Department uses it as part of the assessment of nutrient impacts in mountainous streams (Suplee and Sada de Suplee, 2011). Relative to the control period (1994-1997), HBI scores were essentially unaffected by the nutrient concentration changes observed. Based on the data, it is likely that the streams were N limited in the post-road period and P limited in the post-harvest period.

Beyond this ecoregion, applicable studies are essentially the same as described in **Section 3-1** for the Middle Rockies (excluding Chambers et al., 2011). These studies indicate a range of candidate criteria from 20-30 µg TP/L and 300-1,210 µg TN/L. Work by Mebane et al. (2009) in central Idaho has less direct application here, but note that streams where they observed very low ambient TP and TN concentrations (similar in concentration to Northern Rockies reference streams) N and P co-limitation was the norm.

Conclusion

We recommend 25 µg TP/L and 275 µg TN/L for this ecoregion. The scientific literature most specific to this ecoregion (Welch et al., 1989; Chambers et al., 2011; Gravelle et al., 2009a; Gravelle et al., 2009b) suggests criteria ranging from 20-40 µg TP/L and 210-400 µg TN/L. Some consideration was given to the fact that the natural background concentrations in this ecoregion are quite low relative to the range of potential criteria. The concentrations 25 µg TP/L and 275 µg TN/L result in an N:P ratio of 11, which is higher than the regional reference stream ratio (8:1) but still close to the Redfield range where co-limitation by N and P occurs (the 11:1 ratio suggests slight P limitation).

3.3 LEVEL III: CANADIAN ROCKIES (ECOREGION 41)

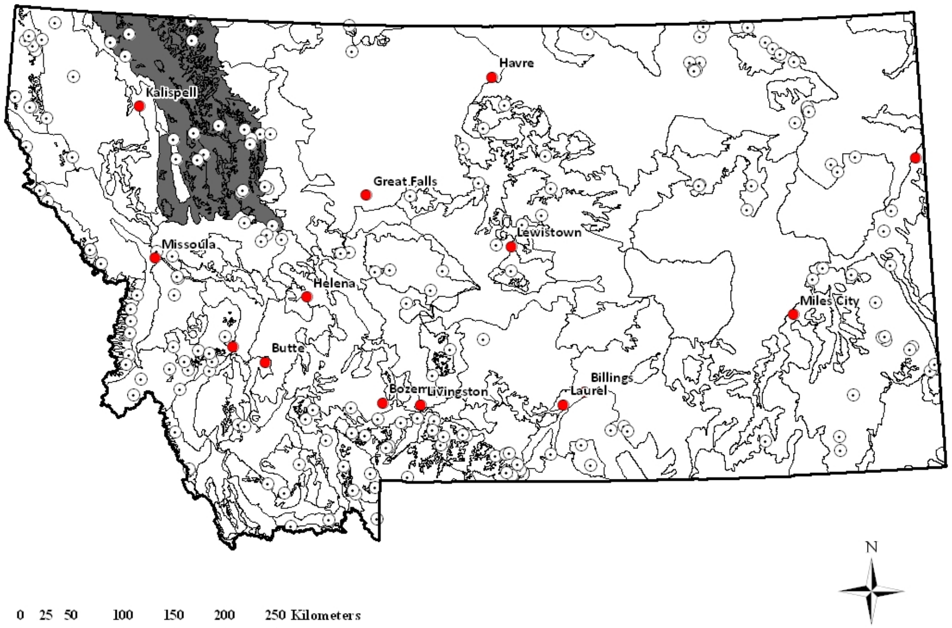

Figure 3-8. Map of Montana showing Canadian Rockies ecoregion.
White dots are the reference sites.

Recommended Numeric Criteria

Total Phosphorus: **25µg TP/L**
Total Nitrogen: **325 µg TN/L**

N:P Ratio of Criteria: **13:1**
N:P Ratio of Reference Sites: **16:1** (Redfield N:P ratio = 7:1)

Descriptive Statistics of Regional Reference Sites

Table 3-4A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Canadian Rockies ecoregion.
Data are from the all-observations dataset after applying Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	13	39	2.5	413	27	63	156	268
TP	14	48	0.5	35	2	4	6	9

Table 3-4B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Canadian Rockies ecoregion.
Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	13	40	80	156	245
TP	14	4	5	6	7

The 25 µg TP/L criterion matches the 97th percentile of reference in the all observations dataset and the 98th percentile of reference in the median dataset. .

The 325 µg TN/L criterion matches the 95th percentile of reference all observations dataset and the 99th percentile of reference in the median dataset.

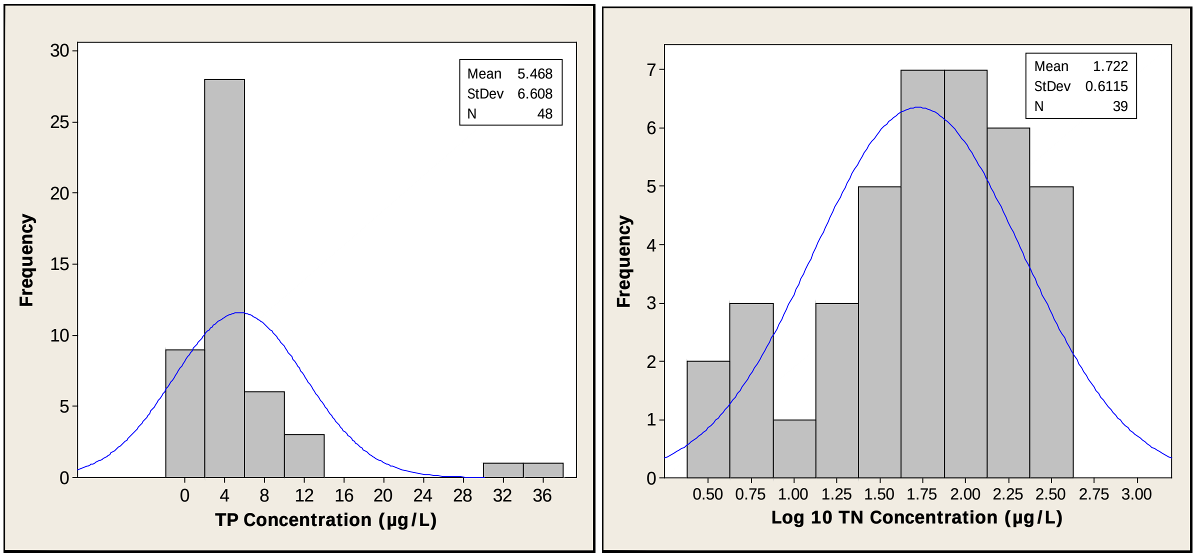

Figure 3-9. Nutrient concentrations from reference streams in the Canadian Rockies ecoregion. Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). Data were collected during the Growing Season (July 1-September 30).

Discussion of the Canadian Rockies Ecoregion Nutrient Criteria

Several studies have direct application to the Canadian Rockies (Sosiak, 2002; Bowman et al., 2007; Scrimgeour and Chambers, 2000). All three were carried out in Canadian rivers in the ecoregion. No model equation between benthic Chl*a* and nutrients was provided in Bowman et al. (2007), however Michelle Bowman graciously provided us the data for the relationship between TP and benthic Chl*a* from their study (personal communication, January 21, 2009). The TP-Chl*a* correlation, though weak, suggests that benthic algal biomass of 150 mg Chl*a*/m^2 equates to 89 µg TP/L, and 125 mg Chl*a*/m^2 equates to 66 µg TP/L (see **footnote 4** for information pertaining to the benthic algae levels used here).

Sosiak (2002) provides a Chl*a* vs. total dissolved phosphate (TDP) + NO_{2+3} multiple-regression equation, and a conversion between TDP and TP concentrations for the Bow River (TP is about 2.8 X TDP). He reports that 150 mg Chl*a*/m^2 equates to 18 µg TP/L (equal to 10 µg TP/L @ 125 mg Chl*a*/m^2). But he assumes that nitrate in the river is essentially saturated (conc. = 267 µg NO_{2+3}-N/L). We reset the NO_{2+3} in the equation to a value (50 µg NO_{2+3}-N/L) that is a much more realistic proportion of any foreseeable TN criterion, with the following results: 150 mg Chl*a*/m^2 corresponds to 41 µg TP/L, and 125 mg Chl*a*/m^2 equates to 23 µg TP/L. Scrimgeour and Chambers (2000) note that, in the absence of human influence, the Wapiti-Smokey rivers are probably P limited, but once alterations to the water quality occur due to kraft mill effluent, N and P co-limitation is most common. Finally, watershed managers on the Bow River are recommending 28 µg TP/L in the central Bow River in order to maintain benthic algae ≤ 150 mg Chl*a*/m^2 (Bow River Basin Council, 2008).

Relevant dose-response studies from outside the ecoregion are essentially the same as those described for the Middle Rockies. Work carried out in northern and southern temperate rivers and streams show that nutrient-benthic Chl*a* regressions have breakpoints at 27-62 µg/L for TP and between 367-602 µg/L for TN (Dodds et al., 2006). Stevenson et al. (2006) show in Michigan streams that the likelihood of reaching bottom coverage by *Cladophor*a of 20-40% increases sharply when TP exceed 30 µg TP/L and 1,000 µg TN/L. This level of streambed coverage by *Cladophora* is unacceptable to the Montana public (see Suplee et al., 2009, Table 1). Chambers et al. (2011) derive nutrient criteria for Canadian streams using multiple methods including dose-response relationships between nutrients and algae and macroinvertebrate metrics. They recommend 20 µg TP/L and 210 µg TN/L for the Montane Cordillera, a mountainous region west and north of Montana in British Columbia (in the Northern Rockies ecoregion). Equations relating benthic algal Chl*a* to total nutrients (Dodds et al., 1997, Equation 17; Dodds et al., 2006; Equation 19), respectively), were used to calculate TN levels that would maintain 125 mg Chl*a*/m^2 given a TP of 25 µg/L. These equations result in TN concentrations ranging from 528-821 µg TN/L. But Bowman et al. (2007) state that nutrient-algae relationships in nutrient-poor lotic systems are harder to predict, and specifically note that Dodds' equations under predict benthic algal biomass of oligotrophic rivers such as those found in the Canadian Rockies. As such, Dodds' equations (and the work of Stevenson et al. (2006)) need to be considered cautiously in this ecoregion.

Conclusion

Total P values derived from Sosiak (2002) and Bowman et al. (Bowman et al., 2007) are in the range of 23 to 89 µg TP/L. The Bow River Basin Council (2008) suggests 28 µg TP/L to maintain river benthic algae at the same levels considered here. None of the equations specific to the Canadian Rockies provide a means to easily derive a TN criterion. Given that the reference sites in this ecoregion have a fairly high TN:TP ratio (16:1, **Table 3-4B**; highest of the western ecoregions), and that others have noted the inherent P limitation of the region (Scrimgeour and Chambers, 2000), it would be prudent to establish TP values that maintain this inherent P limitation. **We recommend 25 µg TP/L, and a corresponding TN criterion (giving consideration to Redfield and the region's natural N:P ratio) of 325 µg TN/L.** Criteria in this ratio (13:1) should induce slight P limitation (as is inherent in the ecoregion), but are shifted somewhat toward a Redfield ratio that would result in co-limitation.

A final note. One level IV ecoregion within the Canadian Rockies, the Southern Carbonate Front (41d), had statistically higher total Kjeldahl N (TKN) concentrations in its reference sites compared to reference sites of the rest of the Canadian Rockies (Varghese and Cleland, 2008; Varghese and Cleland, 2009). Total Kjeldahl N is a close surrogate for TN, therefore we investigated whether or not this level IV ecoregion should have a separate TN criterion. The Southern Carbonate Front's median TN concentration (73 µg TN/L) falls midrange of the Canadian Rockies as a whole (63-80 µg TN/L; **Tables 3-4A, B**), and the Southern Carbonate Front's nitrogen levels are not high enough to warrant separate criteria. The criterion recommended for the Canadian Rockies, 325 µg TN/L, matched the 91st percentile of the Southern Carbonate Front's TN and TKN reference distribution. (When only its TN data were considered, 325 µg TN/L matched the 94th.) Thus, 325 µg TN/L is an appropriate criterion for this level IV ecoregion as well since the great majority of its reference data are less than the criterion.

3.4 LEVEL III: IDAHO BATHOLITH (ECOREGION 16)

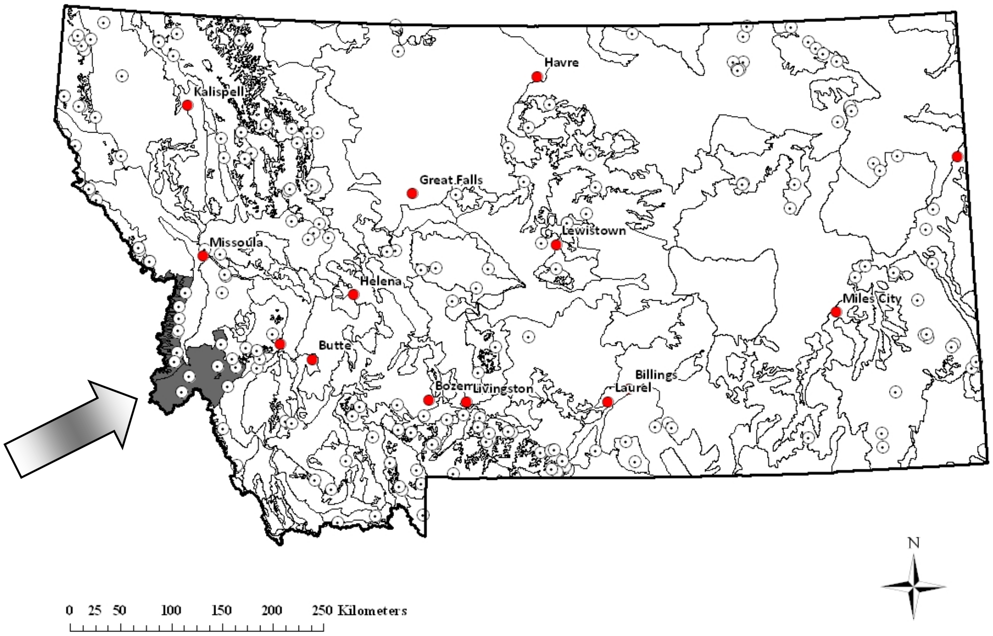

Figure 3-10. Map of Montana showing the Idaho Batholith ecoregion.
White dots are the reference sites.

Recommended Numeric Criteria
Total Phosphorus: **25 µg TP/L**
Total Nitrogen: **275 µg TN/L**

N:P Ratio of Criteria: **11:1**
N:P Ratio of Reference Sites: **10:1** (Redfield N:P ratio = 7:1)

Descriptive Statistics of Regional Reference Sites

Table 3-5A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Idaho Batholith ecoregion.
Data are from the all-observations dataset after applying Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	9	28	2.5	238	46	70	95	163
TP	9	28	0.5	19	4	6	8	11

Table 3-5B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Idaho Batholith ecoregion.
Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	9	40	62	72	104
TP	9	6	6	8	11

The 25 µg TP/L criterion is beyond the 100th percentile of reference for both datasets.
The 275 µg TN/L criterion is beyond the 100th percentile of reference for both datasets.

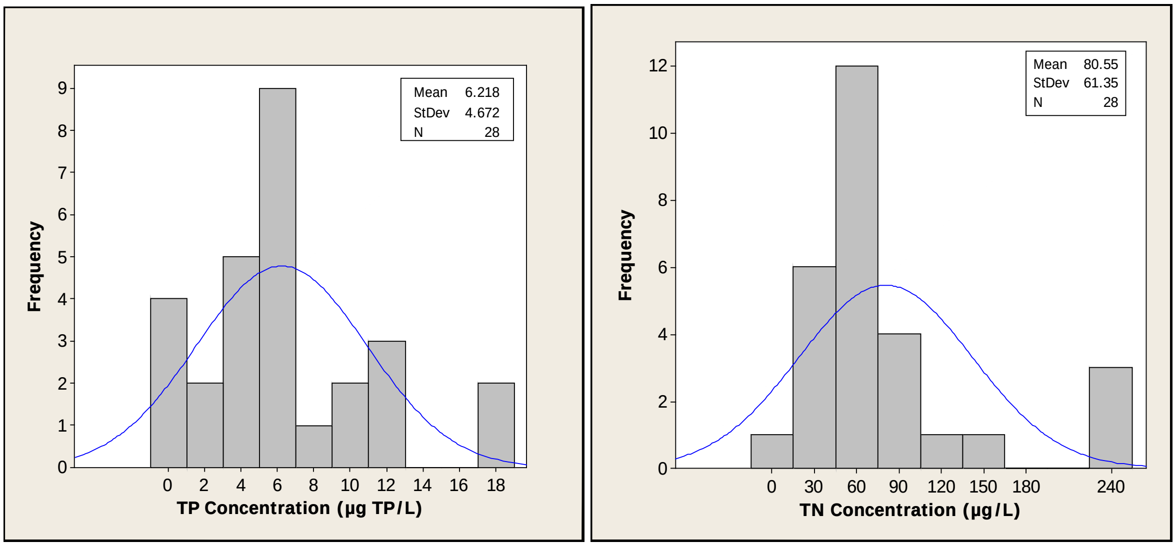

Figure 3-11. Nutrient concentrations from reference streams in the Idaho Batholith ecoregion.
Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). Data were collected during the Growing Season (July 1-September 30).

Discussion of the Idaho Batholith Ecoregion Nutrient Criteria

There is a relatively small extent of this level III ecoregion in Montana (most of it is found in central Idaho). Mebane et al. (2009) carried out a multiple approach, dose-response study in wadeable streams, and several of their sites are located in the Idaho Batholith ecoregion of central Idaho. They carried out *in situ* nutrient limitation trials using nutrient diffusers and other approaches. One of their study sites located in the Idaho Batholith (the Big Wood River) had very low ambient nutrients (7-10 µg TN/L and 50-150 µg TN/L), not unlike reference streams of this ecoregion in Montana (**Tables 3-5A,B; Figure 3-11**). The Big Wood River was found to be very strongly N and P co-limited and, in fact, single N- and P-additions alone grew no more algae than did the un-amended control. This indicates that solo N or P control (no control on the other) would need to be maintaining concentrations no higher than natural background (ca. 8 µg TP/L or 100 µg TN/L) to prevent substantial increases in benthic algal growth. Among the study streams, N and P co-limitation or N-limitation was most common. These data demonstrate that coupled N and P criteria are important to maintain desired water quality conditions. Mebane et al. (2009) recommend criteria of 40 µg TP/L and 600 µg TN/L in order to maintain benthic algae growth to ≤150 mg Chl*a*/m^2, per Suplee et al. (2009). But to maintain 125 mg Chl*a*/m^2, which is more appropriate for many wadeable streams (see **footnote 4**), corresponding concentrations are about 35 µg TP/L and 475 µg TN/L. Note that in Mebane et al.'s study a good response curve was obtained between benthic algal Chl*a* and TN. This was less true for TP where, in a number of cases, there were a fair number of observations falling outside the Chl*a*-TP response curve (i.e., higher benthic algal Chl*a* was observed at lower-than-predicted TP levels). Finally, the study of Mebane et al. (2009) included a number of sites in agricultural settings of the Snake River Plain ecoregion (12), so the study's findings should only be carried so far when applying them to the Idaho Batholith.

Relevant dose-response studies from outside the ecoregion would be the same as those described for the Northern Rockies and Canadian Rockies, which have similar, low nutrient concentrations. Chambers et al. (2011) derive nutrient criteria for Canadian streams using multiple methods, including dose-response of nutrients vs. algae and macroinvertebrate metrics. They recommend 20 µg TP/L and 210 µg TN/L for the Montane Cordillera, a mountainous region in the Northern Rockies ecoregion. Other work in the Northern Rockies suggests values of about 40 µg TP/L and 400 µg TN/L (Gravelle et al., 2009a; Gravelle et al., 2009b). In the Canadian Rockies, a TP criterion could fall between 23 and 89 µg/L and, for the Bow River there, we documented a recommendation of 28 µg TP/L (Bow River Basin Council, 2008). Equations relating benthic algal Chl*a* to total nutrients (Dodds et al., 1997, Equation 17; Dodds et al., 2006; Equation 19) were used to calculate TN levels that would maintain 125 mg Chl*a*/m^2 benthic algae, given a TP of 30 µg TP/L. These equations resulted in TN concentrations from 466-718 µg TN/L. But Bowman et al. (2007) note that nutrient-algae relationships in nutrient-poor lotic systems (like the Idaho Batholith) are harder to predict and, specifically, the equations of Dodds et al. (2006) under predict the actual benthic algal growth observed.

Conclusion

We found that a study that has application to this ecoregion (Mebane et al., 2009) indicates strong N and P co-limitation or N-limitation in the ecoregion's streams, apparently due to very low natural background nutrient concentrations. These low background nutrient concentrations are similar to those observed in the Northern and Canadian Rockies ecoregions (**Table 3-3A, B; Tables 3-4A, B**). The work of Mebane et al. (2009) suggests that good control of nitrogen is probably of greater importance in controlling algal biomass than phosphorus. Bowman et al. (2007) state that nutrient-algae relationships in nutrient-poor lotic systems are hard to predict, and that published models (such as Dodds et al., 2006) under predict actual benthic algal biomass. We gave this finding careful consideration when we evaluated the range of potential criteria for this ecoregion. The work of Mebane et al. (2009) suggests values of 35-40 µg TP/L and 475-600 µg TN/L. The broader range of recommended values that are most applicable to the Idaho Batholith are 20-89 µg TP/L and 210-400 µg TN/L. Consideration was given to the fact that the natural background concentrations in this ecoregion are quite low relative to the range of potential criteria. **We recommend 25 µg TP/L and 275 µg TN/L.** The TN and TP criteria we recommend are in a ratio (10:1) which matches the natural background for the Idaho Batholith (11:1; **Table 3-5B**) and which (based on Redfield) would result in slight P limitation or possibly N and P co-limitation (the latter of which is typically for regional streams).

3.5 LEVEL III: NORTHWESTERN GLACIATED PLAINS (ECOREGION 42)

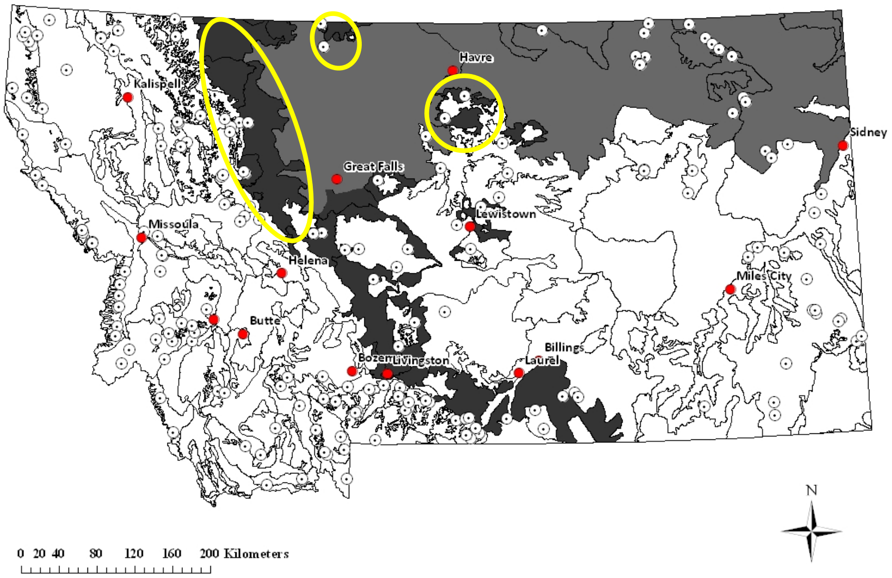

Figure 3-12. Map of Montana showing the Northwestern Glaciated Plains ecoregion (42) in light gray. The dark gray area is a mountain-to-plains transitional zone comprised of level IV ecoregions within ecoregion 42 (and 43, to the south). Mountain-to-plains transitional level IVs that are part of the Northwestern Glaciated Plains (circled in yellow) were not included among the reference data compiled here. White dots are the reference sites.

Recommended Numeric Criteria

Total Phosphorus: **110 µg TP/L**
Total Nitrogen: **1,300 µg TN/L**

N:P Ratio of Criteria: **12:1**
N:P Ratio of Reference Sites: **18:1** (Redfield N:P ratio = 7:1)

Descriptive Statistics of Regional Reference Sites

Table 3-6A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northwestern Glaciated Plains ecoregion.
Data are from the all-observations dataset after applying Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	17	52	55	3891	630	969	1398	1945
TP	18	59	10	638	28	60	111	184

Table 3-6B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northwestern Glaciated Plains ecoregion.
Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	17	720	900	1325	2078
TP	18	26	49	89	105

The 110 µg TP/L criterion matches to the 75th percentile of reference in the all-observations dataset and to the 92nd percentile in the median dataset.

The 1,300 µg TN/L criterion matches to the 65th percentile of reference in the all-observations dataset and to the 72nd percentile in the median dataset.

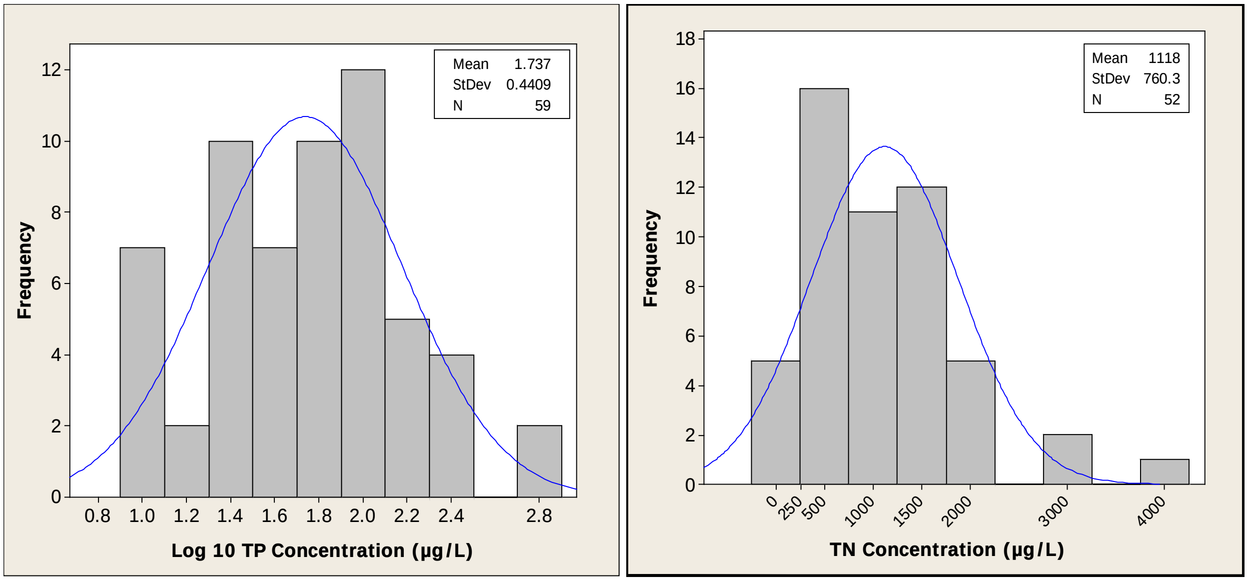

Figure 3-13. Nutrient concentrations from reference streams in the Northwestern Glaciated Plains ecoregion, but excluding data from the level IV ecoregions 42l, 42n, 42q, and 42r.
Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). Data were collected during the Growing Season (June 16-September 30).

Discussion of the Northwestern Glaciated Plains Nutrient Criteria

One study has direct application to the Northwestern Glaciated Plains (Suplee et al., 2008), and there are three additional studies carried out in this ecoregion or in glaciated plains regions further east that have relevance (Wang et al., 2007; Heiskary et al., 2010; Chambers et al., 2011). In Appendix A of Suplee et al. (2008), a relationship is shown between TN concentrations and dissolved oxygen (DO) concentrations as inferred by diatom taxa. Once TN was higher than 1,120 µg TN/L, streams sites had DO concentrations lower than the B-2 DO standard (5.0 mg/L; DEQ, 2010); B-2 streams are widespread throughout this ecoregion. Note in **Tables 3-6A and 3-6B** above that this threshold concentration (1,120 µg TN/L) matches around the 60th percentile of reference; rather low in the reference distribution given that reference sites—by definition— are minimally impacted and support their uses. On close observation of Figure 3.2 in Appendix A of (Suplee et al., 2008), it appears that a concentration between 1,100 and 1,450 µg TN/L could be appropriate as a threshold. Indeed, the 90% confidence interval around the threshold is 780 to 1,480 mg TN/L (Suplee et al., 2008). Giving consideration to the ecoregion's reference distribution (**Table 3-6B; median dataset**), where the 75th percentile of reference equals 1,325 µg TN/L, a TN criterion of about 1,300 µg TN/L appears to be appropriate. It should also be noted that N limitation was strongly indicated in wadeable streams of this region (Suplee, 2004) and was also given consideration in selecting 1,300 µg TN/L. Although the TN:TP ratio of the reference sites (18:1) might suggest the contrary, Redfield ratios are only meaningful when nutrient concentrations are low and generally below saturation (i.e., at concentrations much lower than the natural concentrations found in this ecoregion).

Chambers et al. (2011) derive nutrient criteria for prairie streams of the Northwestern Glaciated Plains ecoregion in Alberta, Canada. They use modeling to relate % agriculture in the watershed and stream nutrient concentrations. In that method, the y-intercept (i.e., zero agriculture) is used to define 'no impact' (Dodds and Oakes, 2004) and can help define a candidate criterion. For the Northwestern Glaciated Plains in Canada this equals 680 µg TN/L. They also use methods involving fixed percentiles of regional reference (and non-reference) sites. Based on all these methods, they provide a range of TN from 680 to 1,110 µg/L, and recommend 980 µg TN/L as a provisional threshold for prairie stream protection. Using the same methods, they also recommend a provisional TP criterion of 106 µg /L.

Fish biometrics have been developed for warm-water plains streams of Montana (Bramblett et al., 2005). The work was carried out exclusively in Montana, in the Northwest Glaciated Plains and the Northwestern Great Plains ecoregion to the south. Fish biometrics provides a means to address harm-to-use for warm-water fish assemblages. One metric, 'proportion of tolerant individuals'[6], shows significant positively correlation with TKN concentrations and significant negative correlation with DO concentrations. Though no specific TN threshold can be drawn from Bramblett et al. (2005), their findings lend support to our work which shows that elevated TN concentrations impact regional DO (and, in turn, the more sensitive taxa in the warm-water fish assemblage).

Patterns in warm-water plains streams in Wisconsin may be roughly comparable to warm-water streams of the Northwestern Glaciated Plains. Wang et al. (2007) examine Wisconsin warm-water streams (e.g., the Southeast Wisconsin Till Plains ecoregion) and relationships between nutrient concentrations and various warm-water fish metrics. One of the metrics ('% individuals considered intolerant')[7] was significantly correlated (negatively) with TP and TN. (This metric is essentially the mirror image of 'proportion of tolerant individuals' of Bramblett et al. (2005).) Wang et al. (2007) also report changepoint thresholds between nutrients and the metric. Nutrient changepoint thresholds represent concentrations above which warm-water fish assemblages are likely to be substantially degraded (Wang et al., 2007). The '% individuals considered intolerant' metric was found to have an ecological threshold at 70-90 µg TP/L and 540 to 1,830 µg TN/L.

Heiskary et al. (2010) recommend numeric nutrient criteria for rivers in different regions of Minnesota, including the southern region of the state which is warm-water and dominated by the Western Corn Belt Plains ecoregion. They examine relationships between DO concentration and fish metrics (including changepoint thresholds), benthic and phytoplankton algae vs. nutrient concentrations, and DO flux (daily maximum minus the daily minimum) vs. invertebrate and fish metrics. Fish metrics include the '% sensitive fish species' metric, which is comprised of most of the same species used in Wisconsin (Lyons, 1992) and suggests there is a fairly consistent bioassessment approach across that plains region. Heiskary et al. (2010) recommend a TP criterion of 150 µg TP/L to protect fish and aquatic life in the southern region of Minnesota.

---

[6] The metric is comprised of highly tolerant species, including: goldfish, common carp, fathead minnow, white sucker, black bullhead, and green sunfish (Bramblett et al., 2005).

[7] This metric was specifically designed to assess fish assemblages in perennial warm-water streams of intermediate size (ie., wadeables), and has direct application to southeastern Minnesota as well (Lyons, 1992). The fish comprising the intolerant species list are almost all warm-water species, and a few are even found in this region of Montana (e.g., smallmouth bass, Iowa darter, and silvery minnow; Brown, 1971).

Earlier we presented work showing that in wadeable streams TP is saturated at 24-62 µg TP/L and TN between 367-602 µg/L (Dodds et al., 2006; Suplee et al., 2012); our recommended criteria for western MT generally are set within or below those ranges since nutrient levels must be below saturation breakpoints to achieve improvements in algae levels/eutrophication. Natural background concentrations of nutrients in the Northwestern Glaciated Plains ecoregion are already at or above these concentrations (Chambers et al., 2011) (see also **Table 3-6A and Table 3-6B**), and yet harm-to-use thresholds at even higher nutrient concentrations can still be identified. If benthic algae are nutrient-saturated, how is this so? We believe it is strongly related to the basic ecology of these streams. The region's streams tend towards two scenario endpoints: (1) un-scoured streams dominated by macrophytes and benthic algae, and (2) scoured, more turbid streams where phytoplankton can be dominant (Suplee, 2004; Suplee et al., 2008, Appendix A). In scenario 2 streams, summer scouring events give phytoplankton a competitive edge because they can tolerate higher turbidity by rotating through the water column via wind and flow advection. Higher turbidity surely induces more light limitation, yet summer phytoplankton concentrations can in cases become very high (>70 µg Chl*a*/L; (Suplee, 2004)). Other regional streams have phytoplankton Chl*a* concentrations as high as 516 µg/L (Suplee, 2004). Phytoplankton-dominated wadeable streams are rarely seen in western MT, and were not a meaningful proportion of the datasets used to derive the nutrient saturation levels mentioned above. In the clearer, un-scoured streams (scenario 1) of the Northwestern Glaciated Plains, macrophytes can—if stimulated enough by nutrients—impact DO concentrations, especially when they senesce (Jewell, 1971). Macrophytes can gain nutrients from the water but also from the sediments (Chambers et al., 1989), obscuring the direct water column nutrient concentration vs. DO relationship. Dissolved oxygen problems probably need to become fairly severe before notable impacts to plains fishes occur, because these fish are already naturally tolerant (Bramblett et al., 2005).

So what we are dealing with here are streams (and associated flora and fauna) of a very different (and more low-DO tolerant) nature relative to western MT, which leads to the higher nutrient impact thresholds observed. That is, to discern a harm-to-use impact in plains streams, levels of nutrients well above the saturation point for algae in non-plains streams are needed in order to override the physical, floral, and faunal differences (and accompanying confounding factors).

Conclusion

Work specific to Montana indicates a TN criterion of 1,300 µg TN/L would maintain DO concentrations at standards common throughout the ecoregion. Studies carried out in warm-water streams in the plains of Canada, in Wisconsin, and in Minnesota provide a range of values from 540 to 1,830 µg TN/L and 70 to 150 µg TP/L. **We recommend 1,300 µg TN/L and 110 µg TP/L.** The TN criterion was derived from a study carried out specifically in this ecoregion in Montana and falls within the range of potential values located in the literature. The TP criterion (110 µg TP/L) is very close to that recommended by Chambers et al. (2011), equates to the 75th-92nd percentile of the reference distribution for this ecoregion (**Tables 3-6A, B**), and falls within the range of criteria located in the literature.

3.5.1 Transitional Level IV Ecoregions within the Northwestern Glaciated Plains: Sweetgrass Upland (42l), Milk River Pothole Upland (42n), Rocky Mountain Front Foothill Potholes (42q), and Foothill Grassland (42r)

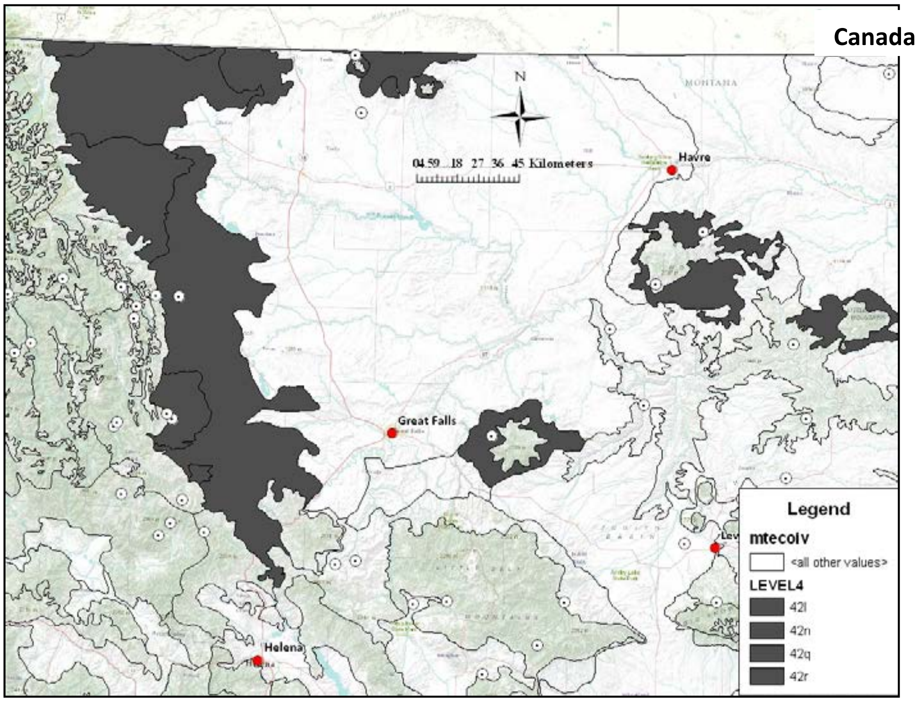

Figure 3-14. Map of Montana showing in gray the transitional level IV ecoregions (42l, 42n, 42q, and 42r) within the Northwestern Glaciated Plains.
White dots are the reference sites.

Recommended Numeric Criteria

Total Phosphorus: **80 µg TP/L**
Total Nitrogen: **560 µg TN/L**

N:P Ratio of Criteria: **7:1**
N:P Ratio of Reference Sites: **10:1** (Redfield N:P ratio = 7:1)

Descriptive Statistics of Regional Reference Sites

Table 3-7A. Descriptive Statistics for TN and TP Concentrations in Transitional Level IV Ecoregions (42q, 42r) of the Northwestern Glaciated Plains. No data were available for 42l, 42n.
Data are from the all-observations dataset after applying Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	5	20	24	2830	115	253	515	704
TP	5	17	1	380	9	20	78	246

Table 3-7B. Descriptive Statistics for TN and TP Concentrations in Transitional Level IV Ecoregions (42q, 42r) of the Northwestern Glaciated Plains.

No data were available for 42l, 42n. Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	5	143	176	200	392
TP	5	11	18	35	206

The 80 µg TP/L criterion matches to the 75th percentile of reference in the all observations dataset and the 79th percentile in the median dataset.

The 560 µg TN/L criterion matches to the 80th percentile of reference in the all observations dataset and is greater than the 100th percentile in the median dataset.

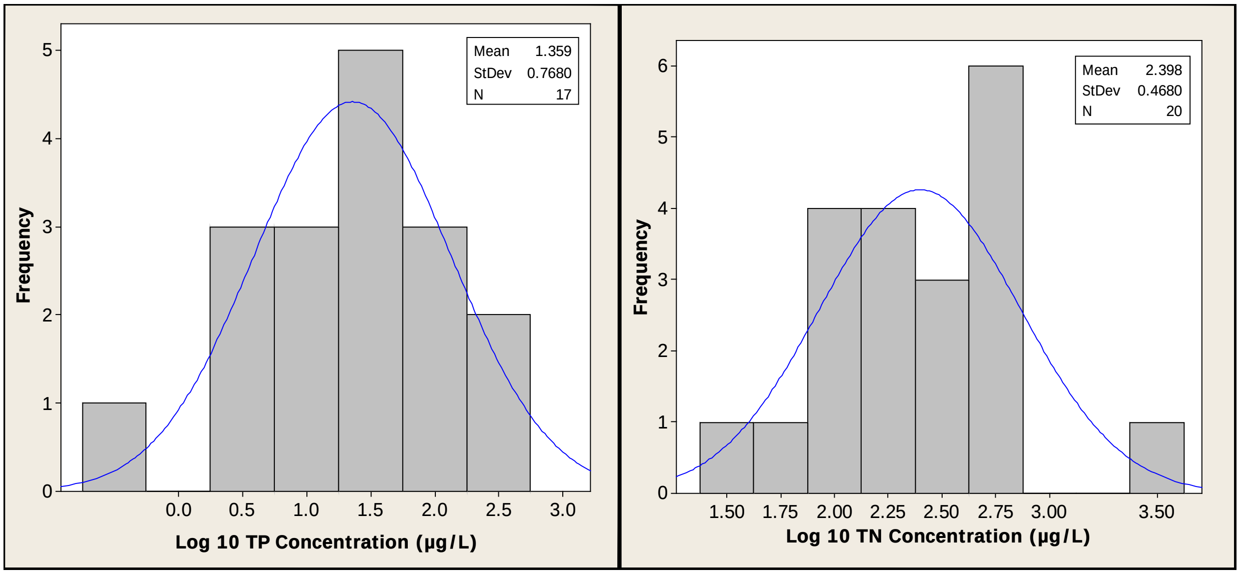

Figure 3-15. Nutrient concentrations from reference streams in the Transitional Level IV ecoregions (42q, and 42r) of the Northwestern Glaciated Plains.

Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). No data were available for 42l and 42n. Data were collected during the Growing Season (June 16-September 30).

Discussion of the Nutrient Criteria for the Transitional Level IV Ecoregions Within the Northwestern Glaciated Plains

In general, streams located in these transitional level-IV ecoregions have more in common with the mountains than the plains. Several lines of information provide support for this. First, although these transitional level IVs form part of the Northwestern Glaciated Plains ecoregion, most of the streams in the transitional level IVs are classified by the state as B-1. This means that they are expected to support salmonid fisheries and are coldwater systems, in sharp contrast to the warm-water streams found further to the east. Clearly, those who developed the stream class system in the late 1950s recognized the strong mountain influences on these streams. Second, floristically they have more in common with mountain streams. Teply and Bahls (2007) carry out a hierarchical cluster analysis on diatom algae (Bacillariophyta) from Montana reference sites, and find that streams of the transitional region are best classified along with the mountain streams. In fact, the level IV ecoregions addressed here (42l, 42n, 42q, and 42r) are being assessed by the Department using diatom metrics designed to evaluate coldwater streams (Teply, 2010; Montana Department Environmental Quality, 2011). Third, although the natural level of nutrients here (**Tables 3-7A, B**) are higher than the most nutrient-rich mountain ecoregion (the Middle Rockies, **Tables 3-1A, B**), they are still much lower than concentrations observed in the Northwestern Glaciated Plains further to the east (**Table 3-6A, B**).

Note that the TP and TN criteria recommended for the four level-III mountain ecoregions to the west (25-30 µg TP/L and 275-325 µg TN/L) fall between the 50th and 60th percentile of reference for these transitional ecoregions (all-observations dataset; **Table 3-7A**). For the median dataset (**Table 3-7B**) the level-III ecoregion criteria ranges fall between the 60th and 67th percentile of reference in this ecoregion for TP, and between the 81st and 85th for TN. Clearly, with these small datasets how the statistical summaries are prepared has a big effect on the distributional statistics reported.

Some reference streams in the transitional ecoregions have benthic algae levels that fall close to or are slightly above the recreationally-based threshold (150 mg Chl*a*/m^2; **Table 3-8**). In reference streams of Montana's mountainous ecoregions (e.g., Middle Rockies, Northern Rockies), we have not measured a site-average benthic Chl*a* level >80 mg/m^2 and the median among sites there is only 14 mg Chl*a*/m^2 (Suplee et al., 2009); also recall that nutrient concentrations of reference streams in the mountain ecoregions are usually well below the recommended criteria. One would expect higher background algae levels in this transitional region given the elevated background nutrient concentrations, and this is what we observe.

Table 3-8. Site-average Benthic Algae Levels Measured in Streams of Ecoregions 42r and 42q.

Site Name	Reference Site No.	Ecoregion (level IV)	Sampling Date	Site-average Chl*a* (mg/m^2)
Clear Creek	ClearCre_121_W	42r	9/20/2001	19
Clear Creek	ClearCre_121_W	42r	8/5/2003	29
Clear Creek	ClearCre_121_W	42r	8/7/2003	17
Clear Creek	ClearCre_121_W	42r	9/7/2003	41
Clear Creek	ClearCre_121_W	42r	8/21/2009	37
Barr Creek lower site at Sun River WMA	BarrCree_504_C	42q	7/9/2009	159
Barr Creek lower site at Sun River WMA	BarrCree_504_C	42q	8/6/2009	84
Barr Creek lower site at Sun River WMA	BarrCree_504_C	42q	9/11/2009	95
Rose Creek upstream from confluence with Barr Creek	RoseCree_518_C	42q	7/10/2009	91
Rose Creek upstream from confluence with Barr Creek	RoseCree_518_C	42q	8/8/2009	148
Rose Creek upstream from confluence with Barr Creek	RoseCree_518_C	42q	9/13/2009	60

WMA – wildlife Management Area

It was necessary to derive benthic algae and nutrient criteria specific to these transitional ecoregions since natural levels of nutrients here are elevated compared to the mountain ecoregions (but still well below concentrations of the plains). Natural algae levels are high enough (i.e., >125mg Chl*a*/m^2) that seasonal DO problems may occur in these streams and may influence the fisheries here. Therefore, the next beneficial use to consider is recreation. Suplee et al. (2009) show that site-average benthic algal Chl*a* levels of 150 mg Chl*a*/m^2 are acceptable to the public, but 200 Chl*a*/m^2 clearly are not. Since benthic algae levels in this region may naturally reach 159 mg Chl*a*/m^2 (**Table 3-8**), but values >200 mg Chl*a*/m^2 have not been observed, a value of 165 mg Chl*a*/m^2 should be an appropriate target. This is about the highest algae level (giving consideration to the statistical confidence around the 150 mg Chl*a*/m^2 average) the public would find acceptable (Suplee and Sada de Suplee, 2011).

Equations relating benthic algal Chl*a* to total nutrients (Dodds et al., 1997, Equation 17; Dodds et al., 2006; Equation 19) were used to calculate TN levels that would maintain 165 mg Chl*a*/m^2 benthic algae based on a TP level of 80 µg/L (80 µg TP/L = 75th-79th percentile of reference here; **Tables 3-7A, B**). These equations resulted in TN concentrations ranging from 445 to 775 µg TN/L. Also, note that nutrient-benthic Chl*a* regressions have saturation breakpoints at 27-62 µg TP/L and between 367-602 µg/L for TN (Dodds et al., 2006), and, the inherent TN:TP ratio of the region is about 13:1 (higher than Redfield).

Conclusion

Giving consideration to the equations of Dodds et al. (2006), saturation thresholds (Dodds et al., 2006), Redfield ratio, and natural background nutrient levels, we **recommend 80 µg TP/L and 560 µg TN/L for the transitional ecoregions 42l, 42n, 42q, and 42r.** The TN criterion is in the range provided by the equations of Dodds et al. (2006), is lower than the higher nitrogen saturation concentration and, along with TP, provides an N:P ratio of 7:1 (at Redfield). Note also that the 560 µg/L TN criterion corresponds to the last commonly-observed concentration in the reference sites (equal to 2.75 log; **Figure 3-15**, right panel).The TP criterion (used to calculate the TN value), having been set to approximately the 75th percentile of regional reference, should prevent unnecessarily high false positive rates (i.e., declaring a reference stream as impaired) when the Department carries out assessments in this region. **The benthic algal biomass criterion for this region is also adjusted up to 165 mg Chl*a*/m^2 to account for natural background levels, with a corresponding Ash Free Dry Mass (AFDM) value equal to 70 g/m^2** (per AFDM, see Suplee et al., 2009, Table 1)**.**

3.6 LEVEL III: NORTHWESTERN GREAT PLAINS (ECOREGION 43) AND THE WYOMING BASIN (ECOREGION 18)

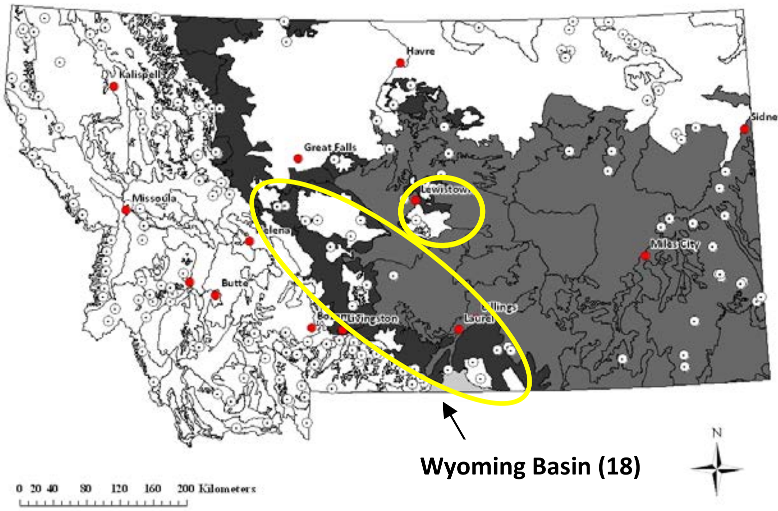

Figure 3-16. Map of Montana showing the Northwestern Great Plains ecoregion (43) in gray, the Wyoming Basin (18) in light gray.

Also, in dark gray, is the mountain-to-plains transitional zone comprising level IV ecoregions in ecoregion 43 (and 42 to the north). Mountain-to-plains transitional level IVs that are part of the Northwestern Great Plains (circled in yellow) were not included among the reference data compiled here. White dots are the reference sites.

Recommended Numeric Criteria

Total Phosphorus: **150 µg TP/L**
Total Nitrogen: **1,300 µg TN/L**

N:P Ratio of Criteria: **9:1**
N:P Ratio of Reference Sites: **13:1** (Redfield N:P ratio = 7:1)

Descriptive Statistics of Regional Reference Sites

Table 3-9A. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northwestern Great Plains ecoregion.

Data are from the all-observations dataset after applying the Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	30	100	50	9900	482	792	1389	3141
TP	32	112	1	9911	36	73	137	519

Table 3-9B. Descriptive Statistics for TN and TP concentrations in Reference Streams of the Northwestern Great Plains ecoregion.

Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	30	537	1042	1365	2296
TP	32	48	82	202	288

The 150 µg TP/L criterion matches to the 77th percentile of reference (all-observations dataset) and the 70th percentile of reference in the median dataset (**but see important caveat in paragraph just above the Conclusion, below**).

The 1,300 µg TN/L criterion matches to the 68th percentile of reference (all observations dataset) and the 73rd percentile of reference in the median dataset.

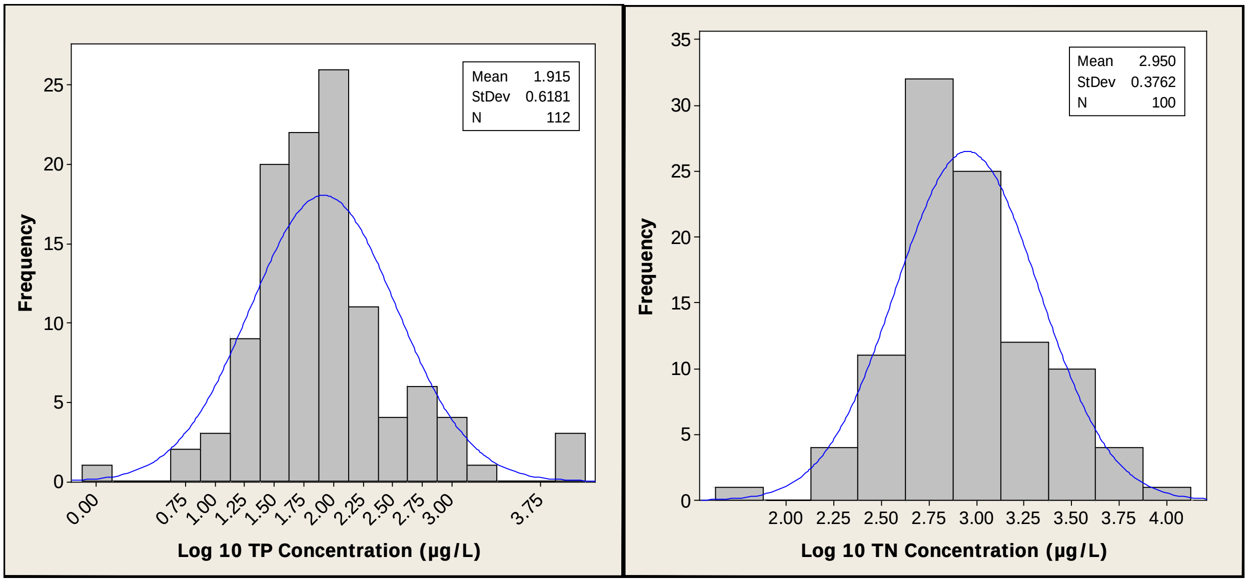

Figure 3-17. Nutrient concentrations from reference streams in the Northwestern Great Plains ecoregion (43), but excluding data from the mountain-to-plains transitional level IV ecoregions 43s, 43t, 43u, and 43v.

Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). Data were collected during the Growing Season (July 1-September 30).

Discussion of the Northwestern Great Plains Nutrient Criteria[8]

The Department (in cooperation with the Carter County Conservation District) carried out a whole-stream nitrogen and phosphorus addition study in Box Elder Creek, a reference stream site located in southeast Montana in the Northwestern Great Plains ecoregion. The full technical report for the study has not yet been prepared, but some of the work has been published (Suplee and Sada de Suplee, 2011). In addition to this study, there are three other studies carried out in plains regions further north and east that should have at least general relevance to this ecoregion (Wang et al., 2007; Heiskary et al., 2010; Chambers et al., 2011).

Appendix B of Suplee and Sada de Suplee (2011) contains key information about the Box Elder Creek dosing study. Additional facts are provided here.

---

[8] The level III ecoregion Wyoming Basin (18) has a very small extent in extreme south central Montana (Figure 3-16). No reference data are available there and for purposes of recommending regional nutrient criteria it is being lumped with the Northwestern Great Plains.

Nutrient dosing took place in summer 2010 and was preceded by "pre" data collection in 2009 and "post" data collected in 2011, at which times no nutrient additions were made. In the study, dissolved sodium nitrate was used as the N source and dissolved dipotassium phosphate as the P source. The High Dose reach was brought up to (after mixing) 150 µg NO_3-N/L and 23 µg SRP/L, continuously, for 53 days in August and September 2010. (Nutrients were added at a single point at the head of the study reach.) Ambient nutrient concentrations in summer are normally 3 µg NO_3-N/L and 4 µg SRP/L (or, in totals, about 500 µg TN/L and 54 µg TP/L). Loading calculations showed that the sodium and potassium added to the stream as part of the compounds used for nutrient additions increased ambient background of those elements by <0.5% and, therefore, are not considered a significant influence on the results. Benthic algae, stimulated by the nutrient additions, grew to levels far above normal for the stream and led to impacts on DO concentrations when the algae senesced *en masse* in early October when the growing season ended. Dissolved oxygen impacts appear to have occurred in patches longitudinally along the stream, with very low DO (zero mg/L) on the bottom in areas where the heaviest densities of decomposing algae settled. The study showed that there is a direct linkage between elevated inorganic nutrients, increased plant growth and, ultimately, impacts to DO standards. Probably the most surprising aspect of the work was that the DO impacts where out-of-phase with peak algal productivity.

Findings from the Box Elder Creek study are generally consistent with the results of the study in Appendix A of Suplee et al. (2008). In that study, DO concentrations—as inferred by diatom taxa—declined when nutrients (TN) became elevated. But what the Box Elder study adds to our understanding is that DO problems may be seasonal and longitudinally patchy in distribution along the stream bottom; in the most impacted locations, DO at the bottom essentially drops to zero. Bramblett et al. (2005), in developing an index of biotic integrity for this region based on fish, find that the 'proportion of tolerant individuals'[9] biometric is significantly correlated (positively) with TKN concentrations and significantly correlated negatively with DO concentrations. We speculate that when nutrient over-enrichment occurs, the more sensitive fish native to these warm-water streams are harmed by patchy, seasonally low DO, and are replaced by tolerant species that can withstand the changes. Indeed, the common carp (*Cyprinus carpio*) is one of the fish in the 'proportion of tolerant individuals' biometric, and is well known for its ability to tolerate very low DO concentrations (Brown, 1971).

Regarding studies from outside the Northwestern Great Plains ecoregion, Chambers et al. (2011) derive nutrient criteria for prairie streams of the Northwestern Glaciated Plains ecoregion in Alberta, Canada. Models that relate % agriculture in the watershed to stream nutrient concentrations are developed, and the y-intercept of the model (i.e., zero agriculture) defines the 'no impact' level (Dodds and Oakes, 2004). For prairie streams in Canada this equals 680 µg TN/L. They also use other methods involving fixed percentiles of regional reference (and non-reference) sites. Taken together, a range of TN concentrations from 680 to 1,110 µg/L is provided, and they recommend 980 µg TN/L as a provisional threshold. Using the same methods, they also recommend a provisional TP criterion of 106 µg /L.

Wang et al. (2007) examine streams in Wisconsin, including warm-water streams located in plains regions of that state (e.g., the Southeast Wisconsin Till Plains ecoregion). Wang et al. (2007) examine relationships between nutrient concentrations and various warm-water fish metrics, one of which ('% individuals considered intolerant') was significantly correlated (negatively) with TP and TN. (This metric is essentially the mirror image of 'proportion of tolerant individuals' of Bramblett et al. (2005). Wang et

---

[9] The metric is comprised of highly tolerant species, including: goldfish, common carp, fathead minnow, white sucker, black bullhead, and green sunfish (Bramblett et al., 2005).

al. (2007) report changepoint thresholds between nutrients and the metric. Nutrient changepoint thresholds represent concentrations above which warm-water fish assemblages are likely to be substantially degraded (Wang et al., 2007). The '% individuals considered intolerant' metric was found to have a threshold between 70 and 90 µg TP/L and 540 to 1,830 µg TN/L (the ranges are due to the different threshold identification techniques used).

Heiskary et al. (2010) recommend numeric nutrient criteria for rivers in different regions of Minnesota, including the southern region of the state which is warm-water and dominated by the Western Corn Belt Plains ecoregion. They examine relationships between DO concentration and fish metrics (including changepoint thresholds), benthic and phytoplankton algae vs. nutrient concentrations, and DO flux (daily maximum minus the daily minimum) vs. invertebrate and fish metrics. Fish metrics include the '% sensitive fish species' metric, which is comprised of most of the same species used in Wisconsin (Lyons, 1992) and suggests there is a fairly consistent bioassessment approach across that plains region. Heiskary et al. (2010) recommend a TP criterion of 150 µg TP/L to protect fish and aquatic life in the southern region of Minnesota.

Total P concentrations in the reference sites at the 75th percentile of reference (**Tables 9A, B**) are near or higher than the highest dose-response TP value we could locate as a potential criterion (150 µg TP/L, Heiskary et al., 2010). This would suggest that applying the 150 µg TP/L value in this region might be futile because the natural levels are already higher. However, the datasets shown include the River Breaks level IV ecoregion for which we are not recommending criteria (more on this next section). It has been standard practice throughout this document to include reference data from break-out level IV ecoregions when describing the overall reference condition for the coarser level III (except where noted for the transitional ecoregions). However we will make an exception here. If the River Breaks data are excluded from the median dataset, the 75th percentile for the Northwestern Great Plains drops to 130 µg/L, considerably lower than the 202 µg/L in **Table 3-9B**. This suggests that in areas outside of the River Breaks the 150 µg TP/L value would be meaningful (and matches the 77th percentile of the Northwestern Great Plains ecoregion not including the River Breaks data).

Conclusion

A scientific study (Box Elder Creek nutrient dosing) carried out in the Northwestern Great Plains in Montana shows a direct linkage between elevated nutrient concentrations and declines in DO concentration. These DO changes likely impact warm-water fish assemblages periodically and may lead to the undesirable changes in local fish assemblages observed by Bramblett et al. (2005). A study carried out in the Northwestern Glaciated Plains of Montana (Suplee et al., 2008, Appendix A) suggests a criterion of 1,300 µg TN/L to maintain DO concentrations at state standards. Studies carried out in warm-water streams in the plains of Canada, in Wisconsin, and in Minnesota provide a range of values from 540 to 1,830 µg TN/L and 70 to 150 µg TP/L. **We recommend 1,400 µg TN/L and 150 µg TP/L for this ecoregion.** The TN criterion for the Northwestern Glaciated Plains study (1,300 µg TN/L) is has good application to this ecoregion. This is supported by the similarity in the central tendency of the two regions' reference TN concentrations (**Tables 3-6B and 3-9B**), and a general similarity in the ecology of the streams in the two regions (see the discussion of plains streams ecology in the last paragraph of **Discussion of the Northwestern Glaciated Plains Nutrient Criteria**). A concentration of 1,300 µg TN/L matches the 68th and 73rd percentiles of reference (all-observations and median datasets, respectively).

The suggested TP criterion (150 µg/L) matches that recommended by Heiskary et al. (2010). **Table 3-9B** shows that natural background is higher than this concentration, but the statistics include the River Breaks level IV which has naturally high TP concentrations and for which we will not be recommending criteria (more on that, next section). **If River Breaks TP data are excluded from the Northwestern Great Plains ecoregion median dataset, the 150 µg TP/L criterion then matches the 77th percentile of reference (which makes the selection of 150 µg TP/L more reasonable).** Additional work and analysis is needed to refine the TP criterion for this ecoregion.

Note: No nutrient criteria are being proposed for the River Breaks level IV ecoregion (discussed next). As such, the criteria recommended above for the Northwestern Great Plains would not apply there.

3.6.1 Level IV Ecoregion within the Northwestern Great Plains: River Breaks (43c)

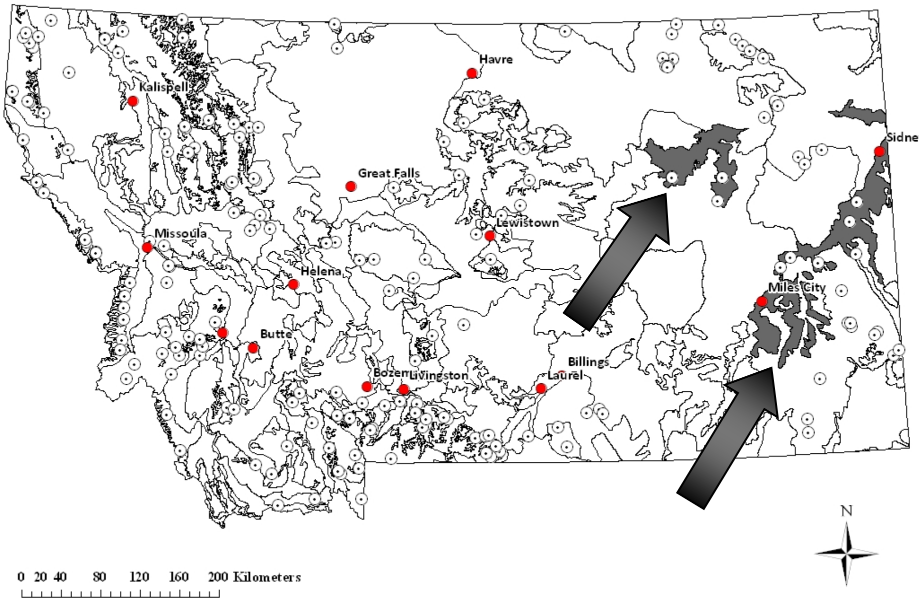

Figure 3-18. Map of Montana showing the River Breaks (43c), a level IV ecoregion within the Northwestern Great Plains ecoregion.
White dots are the reference sites.

Recommended Numeric Criteria

Total Phosphorus: **NONE RECOMMENDED**
Total Nitrogen: **NONE RECOMMENDED**

N:P Ratio of criteria: **n/a**
N:P Ratio of Reference sites: **8:1** (Redfield N:P ratio = 7:1)

Descriptive Statistics of Regional Reference Sites

Table 3-10A. Descriptive Statistics for TN and TP Concentrations in Reference Streams of the River Breaks (43c) level IV ecoregion.
Data are from the all-observations dataset after applying the Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	8	28	480	9900	1005	1333	2486	3792
TP	8	29	33	9911	51	129	293	2123

Table 3-10B. Descriptive Statistics for TN and TP Concentrations in Reference Streams of the River Breaks (43c) level IV Ecoregion.
Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	8	1158	1213	2071	2270
TP	8	78	153	257	301

Table 3-10C. Descriptive Statistics for NO_{2+3} and SRP Concentrations in Reference Streams of the River Breaks (43c) level IV Ecoregion.

Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
Nitrate + nitrite (NO_{2+3})	8	3	7	241	606
Soluble reactive P (SRP)	7	5	6	18	28

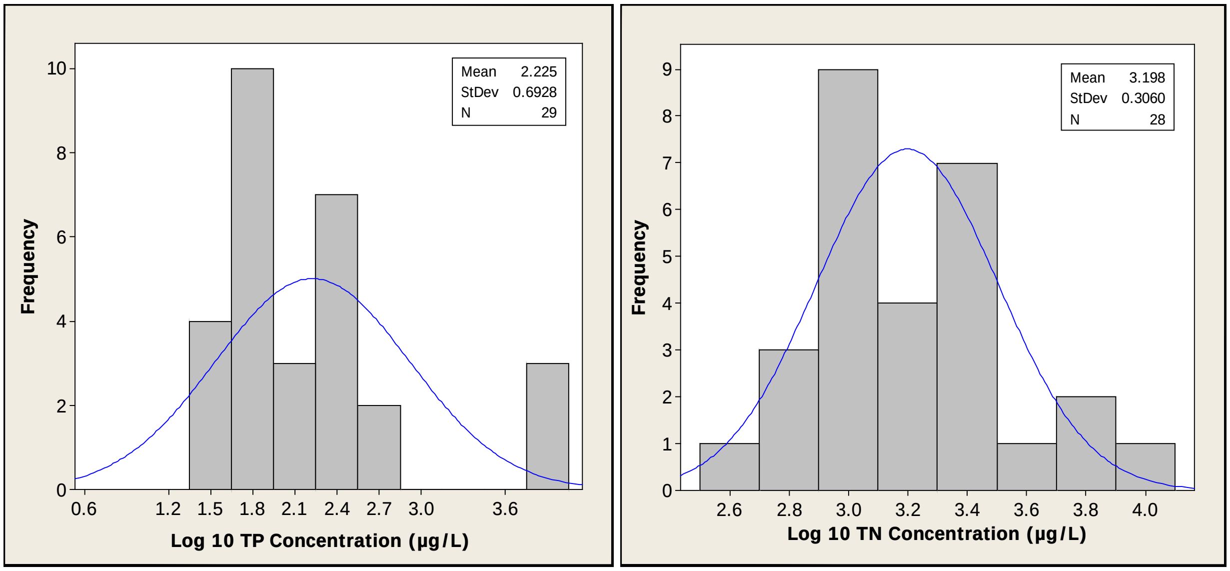

Figure 3-19. Nutrient concentrations from reference streams in the River Breaks (43c) level IV ecoregion.

Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). Data are from the Growing Season (July 1-September 30).

Discussion of the River Breaks (43c) Nutrient Criteria

As shown in **Tables 3-10A** and **B**, average TN concentrations in the reference sites of this level IV ecoregion are near to or higher than harm-to-use levels identified for the Northwestern Great Plains and the Northwestern Glaciated Plains (i.e., 1,300 µg TN/L). On the phosphorus side, the highest dose-response TP criterion identified so far for the plains (150 µg TP/L;(Heiskary et al., 2010) corresponds to the 57th percentile in the all-observations dataset and the 50th percentile in the median dataset, i.e., about average for this ecoregion's reference sites.

In this ecoregion we have also presented the soluble nutrient concentrations (**Table 3-10C**). In relation to algal biomass, it results that soluble nutrients here are saturated or nearly so in the reference sites. For NO_2+NO_3, the 75th percentile of River Breaks reference streams was 241 or 631 µg N/L (median or all-observations datasets, respectively; all observations table not shown). Rier and Stevenson (2006) show there is little peak algal biomass increase above 308 µg soluble nitrogen per liter (and peak biomass may actually be saturated closer to 250 µg DIN/L). As such, River Breaks reference sites are often saturated with soluble nitrogen. Soluble P concentrations are also quite high. At the 75th percentile of reference SRP is 18 or 20 µg P/L (median vs. all-observations datasets, respectively; all observations table not shown) and therefore these reference streams are often P saturated for peak algal biomass, or nearly so (Bothwell, 1989; Horner et al., 1983).

Clearly these streams have highly elevated nutrient levels naturally, but they also have characteristics that strongly dampen plant growth as they have not been found to develop a robust benthic flora. Highly dissected and erodible terraces and uplands lead to bottomlands of this ecoregion where the soils have poor permeability (Woods et al., 2002). This results in flashy, sediment-laden flows when summer thunderstorms occur. All eight of the reference sites in this region have been found to be extremely turbid when sampled in summer (e.g., as high as 30,000 mg TSS/L with accompanying turbidity of 4,000 nephelometric turbidity units), although by fall in some cases the water has cleared as summer thunderstorms diminish.

Benthic algal growth in these streams (due to the factors above) is usually low (**Figure 3-20),** lower than what is observed in streams for the plains regions as a whole (see data pertaining to the plains in the July 16, 2009 presentation to the Nutrient Work Group, *available at*: http://deq.mt.gov/wqinfo/nutrientworkgroup/AgendasMeetingsPresentations.mcpx). Macrophyte density also tends to be limited. In half of the eight reference streams from this ecoregion no macrophytes were observed at all, in two streams they were present but extremely sparse, and in two streams they were commonly found at sparse to moderate levels along much of the stream channel. As for benthic algae, flashy turbid flows and lengthy periods of high turbidity in these streams prevent a robust benthic flora from developing in many cases, in spite of abundant nutrient availability.

As has been observed in streams along Montana's Hi-line (Suplee, 2004), plains streams can at times develop high levels of phytoplankton Chl*a*. In this ecoregion, especially at these nutrient levels, this occurs in the reference sites as well. Many phytoplankton Chl*a* observations from the reference streams in the River Breaks are low (e.g., < 10 µg Chl*a*/L) but in quite a few cases they can become high (e.g., 72 µg Chl*a*/L, Hart Creek, 7/30/2006; 44 µg Chl*a*/L, Snap Creek, 8/24/2006). Suplee (2004) found that 95% of the phytoplankton samples from reference streams in the Northwestern Glaciated Plains were <20 µg Chl*a*/L. In comparison to the River Breaks, this suggests that—at least sometimes—River Breaks reference streams have naturally high phytoplankton Chl*a* concentrations due to ecological conditions (and elevated nutrients) prevalent in the ecoregion. Whether or not these high levels of phytoplankton Chl*a* affect the region's fish fauna is unknown. In all probability, the more severe physical constraints here (i.e., flashy conditions with extreme levels of suspended sediment) are a far greater constraint on the fish fauna and other aquatic life.

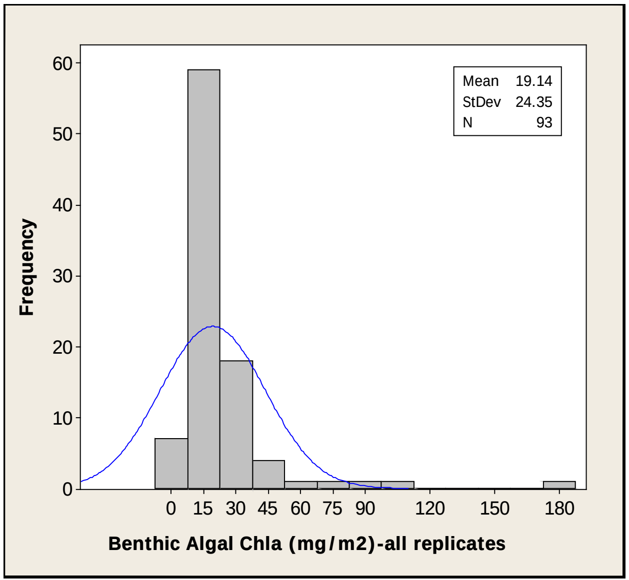

Figure 3-20. Benthic algal density (mg Chl*a*/m^2), all replicates, from reference streams in the River Breaks (43c) ecoregion.
Data are from the Growing Season (July 1-September 30).

Conclusion

This level IV ecoregion has highly turbid, flashy streams with naturally elevated TP and TN levels and soluble nutrient concentrations at or above saturation levels needed to support maximum algal biomass. Concentrations observed in the region's reference sites indicate that nutrient concentrations here are already near to or elevated above the harm-to-use thresholds identified for the plains region as a whole. **As such, no nutrient criteria are recommended for streams within this level IV ecoregion.** Readers should note that the nutrient criteria recommended for the Northwestern Great Plains (level III), discussed previously, would apply across that ecoregion, except here in the River Breaks.

3.6.2 Transitional Level IV Ecoregions within the Northwestern Great Plains: Non-calcareous Foothill Grassland (43s), Shields-smith Valleys (43t), Limy Foothill Grassland (43u), Pryor-Bighorn Foothills (43v), and Parts of the Unglaciated Montana High Plains (43o)[10]

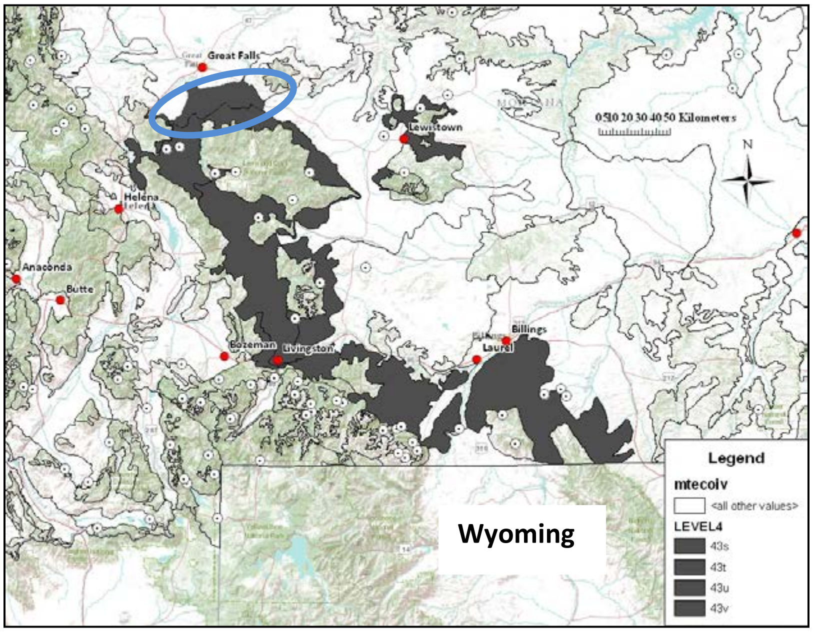

Figure 3-21. Map of Montana showing in gray the transitional level IV ecoregions (43s, 43t, 43u, and 43v) within the Northwestern Great Plains.
The sub-section of ecoregion 43o which is grouped with these level IV ecoregions is located just south of Great Falls (circled in blue). White dots are the reference sites.

Recommended Numeric Criteria
Total Phosphorus: **33 µg TP/L**
Total Nitrogen: **440 µg TN/L**

N:P Ratio of Criteria: **13:1**
N:P Ratio of Reference Sites: **13:1** (Redfield N:P ratio = 7:1)

Table 3-11A. Descriptive Statistics for TN and TP Concentrations in Transitional Level IV Ecoregions (43s, 43t, 43u) of the Northwestern Great Plains.
No data were available for 43o, 43v. Data are from the all-observations dataset after applying the Brillouin Evenness Index.

			Nutrient Concentration (µg/L)					
					Conc. at given Percentile			
Nutrient	**Number of Reference Sites**	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	12	40	50	753	78	112	174	224
TP	12	40	3	108	6	10	22	34

---

[10] For the Unglaciated Montana High Plains ecoregion, only the polygon located just south of Great Falls, MT is associated with this transitional region.

Table 3-11B. Descriptive Statistics for TN and TP Concentrations in Transitional Level IV Ecoregions (43s, 43t, 43u) of the Northwestern Great Plains.
No data were available for 43o, 43v. Data are from the median dataset.

		Concentration at given Percentile (µg/L)			
Nutrient	**Number of Reference Sites**	**25th**	**(Median) 50th**	**75th**	**90th**
TN	12	98	134	202	222
TP	12	8	10	27	33

The 33 µg TP/L criterion matches to the 87th percentile of reference (all observations dataset) and the 90th percentile of reference in the median dataset.

The 440 TN/L criterion matches to the 98th percentile of reference in the all-observations dataset and is greater than the 100th percentile in the median dataset.

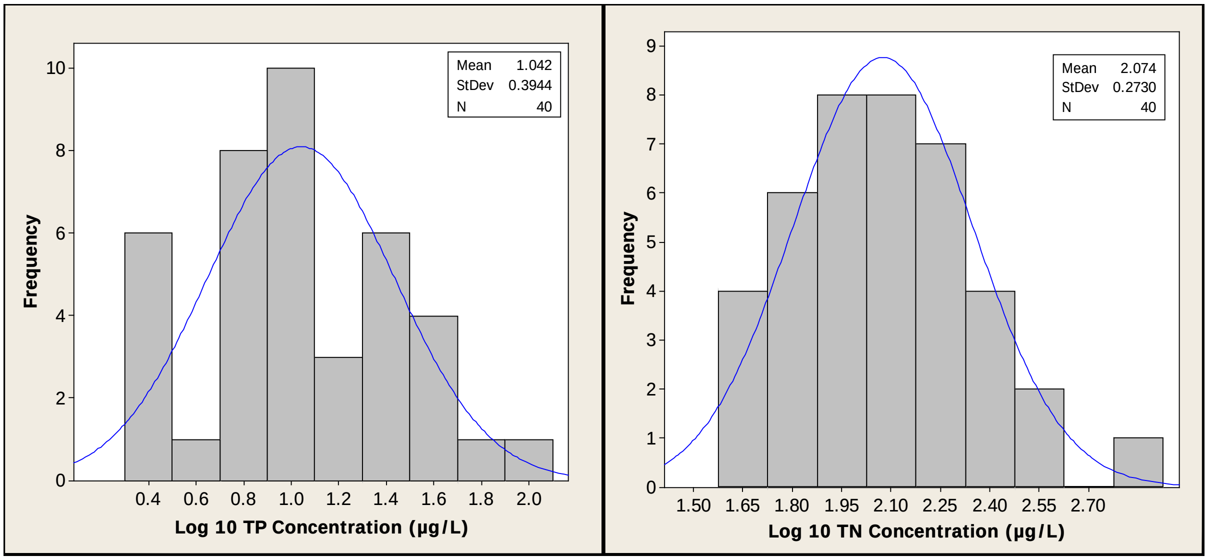

Figure 3-22. Nutrient concentrations from reference streams in the transitional level IV ecoregions (43s, 43t, 43u) of the Northwestern Great Plains.
Data shown are from the all observations dataset (after application of the Brillouin Evenness Index). No data were available for 43o or 43v. Data were collected during the Growing Season (July 1-September 30).

Discussion of the Nutrient Criteria for the Transitional Level IV Ecoregions Within the Northwestern Great Plains

In general, streams located in these transitional level-IV ecoregions have more in common with the mountains than the plains. Several lines of information support this. First, although these transitional level IVs form part of the Northwestern Great Plains ecoregion, virtually all streams in the transitional level IVs are classified by the state as B-1. This means that they are expected to support salmonid fisheries and are generally coldwater systems, in sharp contrast to the warm-water streams found in the Northwestern Great Plains further to the east. It is clear that when the state's stream-class system was developed in the late 1950s, its developers recognized the strong mountain influences on these streams. Second, floristically they have more in common with mountain streams. Teply and Bahls (2007) carried out a hierarchical cluster analysis on diatom algae (Bacillariophyta) from Montana reference sites, and find that streams of the transitional region are best classified along with the mountain streams. In fact, level IV ecoregions addressed here (43s, 43t, 43u, 43v and part of 43o) are being assessed by the Department using diatom metrics for coldwater streams (Teply, 2010; Montana Department

Environmental Quality, 2011). Third, the natural levels of nutrients here (**Tables 3-11A, B**) are not unlike the most nutrient-rich mountain ecoregion (the Middle Rockies, **Tables 3-1A, B**), but they are much lower than the Northwestern Great Plains further to the east (**Tables 3-9A, B**).

In contrast to nutrient concentrations, which are similar to the Middle Rockies, site-average benthic algae levels in these transitional region's reference sites are somewhat higher than the Middle Rockies (**Figure 3-23**). Of specific interest is a high average value (170 mg Chl*a*/m^2) from the Elk Creek reference site (reference site No. ElkCreek_511_C; ecoregion 43u). The high benthic algae density likely resulted from the stream's naturally elevated TP, where summer concentrations ranged from 29 to 41 µg/L (average: 31 µg TP/L). However, the data do not suggest that elevated nutrients are a common factor across the level IV ecoregion in which Elk Creek resides (Limy Foothill Grasslands, 43u). Another reference site there (Middle Fork Judith River; MFJudith_513_C) has TN and TP concentrations well below the median of the aggregate reference sites shown in **Table 3-11A**.

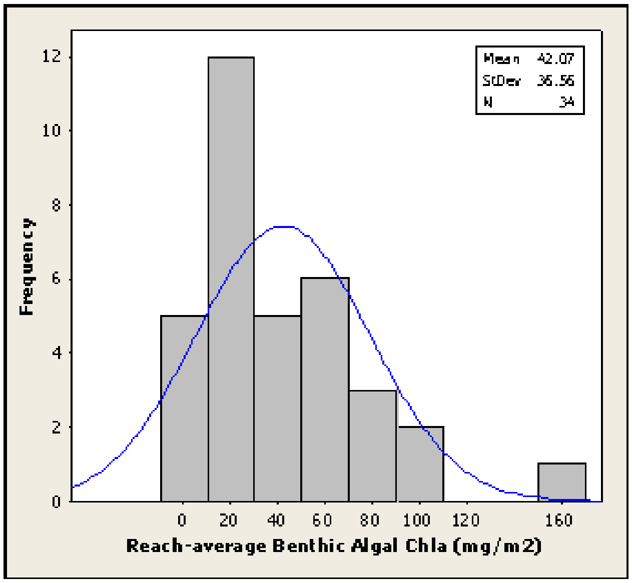

Figure 3-23. Site-average[11] benthic algae density (mg Chl*a*/m^2) from reference streams in the transitional level IV ecoregions (43s, 43t, 43u) of the Northwestern Great Plains ecoregion.

The transitional ecoregions of the Northwestern Great Plains (43s, 43t, 43u) have natural nutrient concentrations with a central tendency roughly comparable to the Middle Rockies. Further, most reference streams (9 of 10) of 43s, 43t, and 43u have benthic algae levels lower than the thresholds considered throughout this document (125 and 150 mg Chl*a*/m^2). However there is one exception to this, Elk Creek, where benthic algae was above the thresholds. This finding is in contrast to the transitional ecoregions of the Northwestern *Glaciated* Plains (see **Section 3.5.1**), where 2 of 3 reference sites had benthic algae levels above the thresholds; there, higher algae levels (and nutrients) seem to be the norm. Because elevated algae levels seem to be the exception and not the rule here in the transitional ecoregions of the Northwestern Great Plains, we used the 125 and 150 mg Chl*a*/m^2 thresholds to help derive the nutrient criteria. We also gave consideration to the naturally higher TP

---

[11] A site average is the arithmetic mean of (normally) 11 sample replicates collected along a short stream reach ≥ 150 m in length using an unbiased systematic approach. See DEQ Standard Operating Procedure manual for benthic Chl*a* (Montana Department of Environmental Quality, 2011).

observed in the Elk Creek site. Elk Creek's average summer TP matches the 85th percentile of the aggregate reference distribution (**Table 3-11A**).

Conclusion

Equations relating benthic algal Chl*a* to total nutrients (Dodds et al., 1997, Equation 17; Dodds et al., 2006; Equation 19) were used to calculate TN levels that would maintain 125 and 150 mg Chl*a*/m^2 benthic algae, respectively, based on a TP concentration of 33 µg/L (about the 90th percentile of reference [**Tables 3-11A, B**] and a bit higher than the average concentration observed in the Elk Creek reference site). These equations resulted in TN concentrations ranging from 439 to 1,125 µg TN/L. Also, we bore in mind the fact that nutrient-benthic Chl*a* regressions have saturation breakpoints at 27-62 µg TP/L and between 367-602 µg/L for TN (Dodds et al., 2006). And, as for other ecoregions, we took into account the TN:TP ratio of the region, which is about 13:1 (suggests slight P limitation of benthic algae). **We recommend 33 µg TP/L and 440 µg TN/L as criteria for the transitional ecoregions 43s, 43t, 43u, 43v, and part of 43o**. The TN criterion is within the range provided by the equations of Dodds et al. (2006), and within the range of saturation thresholds provided by the same authors. Together, the TN and TP criteria provide an N:P ratio of 13:1 (which matches the ratio in the regions reference sites).

4.0 Reach-specific Numeric Nutrient Criteria Recommendations

In **Section 3.0**, ecoregions were used as the ecologically-based system for segregating nutrient criteria for different geographic zones. However, The Department recognizes that within each ecoregional zone there are streams with unique characteristics where numeric nutrient criteria must be considered on a case-by-case basis. Conditions that could render the ecoregional criteria inappropriate include, for example, the presence of a large dam-regulated lake or reservoir upstream[12], or the upstream influence of a level-IV ecoregion known to have elevated TP concentrations. A few cases have already been identified, and these are presented here. Readers should note that outside of the specified reaches described here, the ecoregion-wide criteria would apply. The Department recognizes that other reach-specific exceptions to the ecoregional criteria may be identified in the future; these can be addressed on a case-by-case basis going forward.

4.1 Flint Creek: From the Georgetown Lake Outlet to The Boundary of Ecoregion 17ak at Latitude 46.4002, Longitude -113.3055

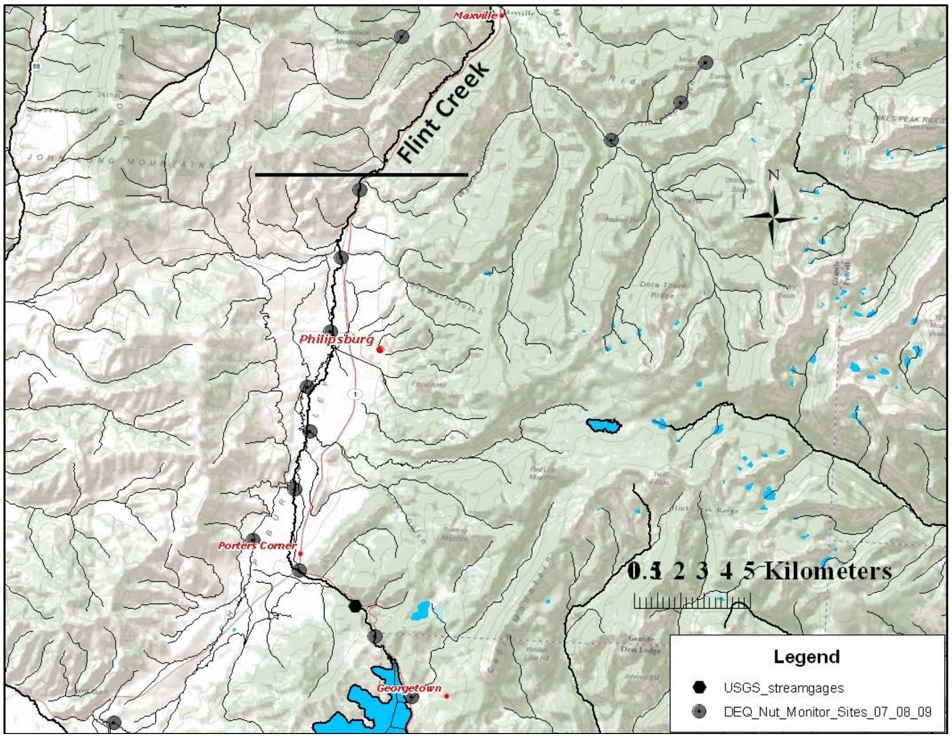

Figure 4-1. Map showing the Flint Creek watershed below the Georgetown Lake outlet.
The criteria presented here would apply from the lake outlet downstream to the black horizontal line, which is the boundary of ecoregion 17ak (Deer Lodge-Philipsburg-Avon Grassy Intermontane Hills and Valleys).

---

[12] When it comes to reservoirs and dam-regulated lakes, specific state laws must be considered. Conditions resulting from the reasonable operation of dams on July 1, 1971 are natural (§75-5-306[2], MCA). There may exist reasonably operated dams that, due to the nature of the water releases, characteristics of the reservoir, etc., result in nutrient concentrations (and possibly benthic algal densities) that are higher than the ecoregionally-based criteria recommended. These situations will generally be considered by the Department on a case-by-case basis.

Recommended Numeric Criteria

Total Phosphorus: **72 µg TP/L**
Total Nitrogen: **500 µg TN/L**

N:P Ratio of criteria: **7:1**
N:P Ratio of Flint Creek's Water Source: **9:1** (Redfield N:P ratio = 7:1)

Table 4-1. Descriptive Statistics for TN and TP concentrations in Flint Creek just Below Georgetown Lake Outlet (July through September).

		Nutrient Concentration (µg/L)					
				Conc. at given Percentile			
Nutrient	**Number of Samples**	**Min**	**Max**	**25th**	**(Median)50th**	**75th**	**90th**
TN	15	75	1200	239	340	419	585
TP	18	5.0	161	19	36	72	99

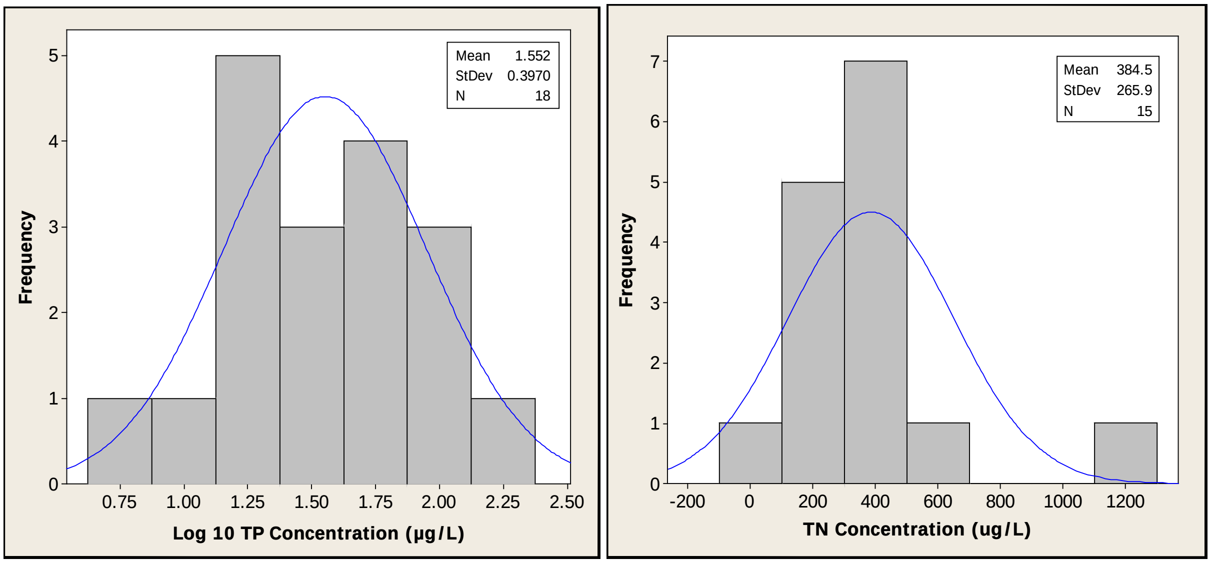

Figure 4-2. Nutrient concentrations observed in Flint Creek just downstream of the point where water exits Georgetown Lake through the dam.
Data were collected during the Growing Season (July 1-September 30).

Discussion of the upper Flint Creek Nutrient Criteria

Assessments of Georgetown Lake by the Department in the late 1990s indicated that the lake was fully supporting its beneficial uses. The lake has high levels of internal nutrient loading, particularly for phosphorus, and nutrient concentrations in the hypolimnion can become quite elevated especially during periods when hypolimnetic dissolved oxygen becomes low (Mortimer, 1941). The lake's dam forms the headwaters of Flint Creek and the stream receives water from the lake through a 36 inch diameter pipe. Flow into the pipe is controlled by a slide gate at the upstream end of the pipe on the dam. Data collected by the Department indicates that, at times, the intake intersects the lake's hypolimnion and can introduce elevated nutrients to the stream below. State law requires that reasonable dam operations be considered natural. Therefore we consider the dam operation effects when we developed nutrient criteria for upper Flint Creek.

Table 4-1 and **Figure 4-2** show nutrient concentrations measured in Flint Creek very near to where water comes out of Georgetown dam. These data (collected by the Department and others) were collected between 2005 and 2009 during the July-September period. The data show that the ambient concentrations coming out of the lake are, during the summer when nutrient criteria apply, elevated compared to natural background for the Middle Rockies (**Tables 3-1A, B**). The extent to which these conditions persist downstream was also evaluated. This was difficult to determine precisely, but the data suggest that by Flint Creek station 9 the stream has returned to "normal" Middle Rockies nutrient concentrations. (Station 9 is the gray dot immediately south of the horizontal line dividing Flint Creek in **Figure 4-1.**) Coincidently, station 9 is very near the boundary of ecoregion 17ak so, for ease, we have set the termination point of Flint Creek's reach-specific criteria there.

Equations relating benthic algal Chl*a* to total nutrients (Dodds et al., 1997, Equation 17; Dodds et al., 2006; Equation 19) were used to calculate TN levels that would maintain 125 and 150 mg Chl*a*/m^2 benthic algae, respectively, based on a TP concentration of 72 µg/L (equal to the 75th percentile of the data for Flint Creek just below Georgetown Lake). These equations resulted in TN concentrations ranging from 290 to 637 µg TN/L (290-380 µg TN/L @ 125 mg Chl*a*/m^2, and 394-637 µg TN/L @ 150 mg Chl*a*/m^2). Also, we considered that nutrient-benthic Chl*a* regressions have breakpoints at 27-62 µg TP/L and between 367-602 µg/L for TN (Dodds et al., 2006). (Because of the location of the water out take from Georgetown Lake, Flint Creek starts with TP concentrations already above saturation.) And, we took into account the TN:TP ratio of water in Flint Creek just below the dam, which is about 9:1 (still in the co-limitation range).

Conclusion

We recommend 72 µg TP/L and 500 µg TN/L as criteria for the reach of Flint Creek between Georgetown Lake dam and the boundary of ecoregion 17ak, which is located at 46.4002 latitude, -113.3055 longitude. The TN criterion matches the 88th percentile of the water-quality data coming out of the dam into Flint Creek. The TN concentrations calculated for 125 mg Chl*a*/m^2 could not be consistently achieved without requiring that the Georgetown Lake outtake be raised above the level of the hypolimnion (**Table 4-1**). 500 µg TN/L is within the range provided by the equations of Dodds et al. (2006) for 150 mg Chl*a*/m^2, and within the range of saturation thresholds provided by the same authors. Together, the TN and TP criteria provide an N:P ratio of 7:1 (at Redfield). **The benthic algal biomass criterion for this region is set at 150 mg Chl*a*/m^2 to account for the dam-elevated nutrient levels, with a corresponding AFDM value equal to 45 g/m^2.**

4.2 BOZEMAN CREEK, HYALITE CREEK, AND EAST GALLATIN RIVER

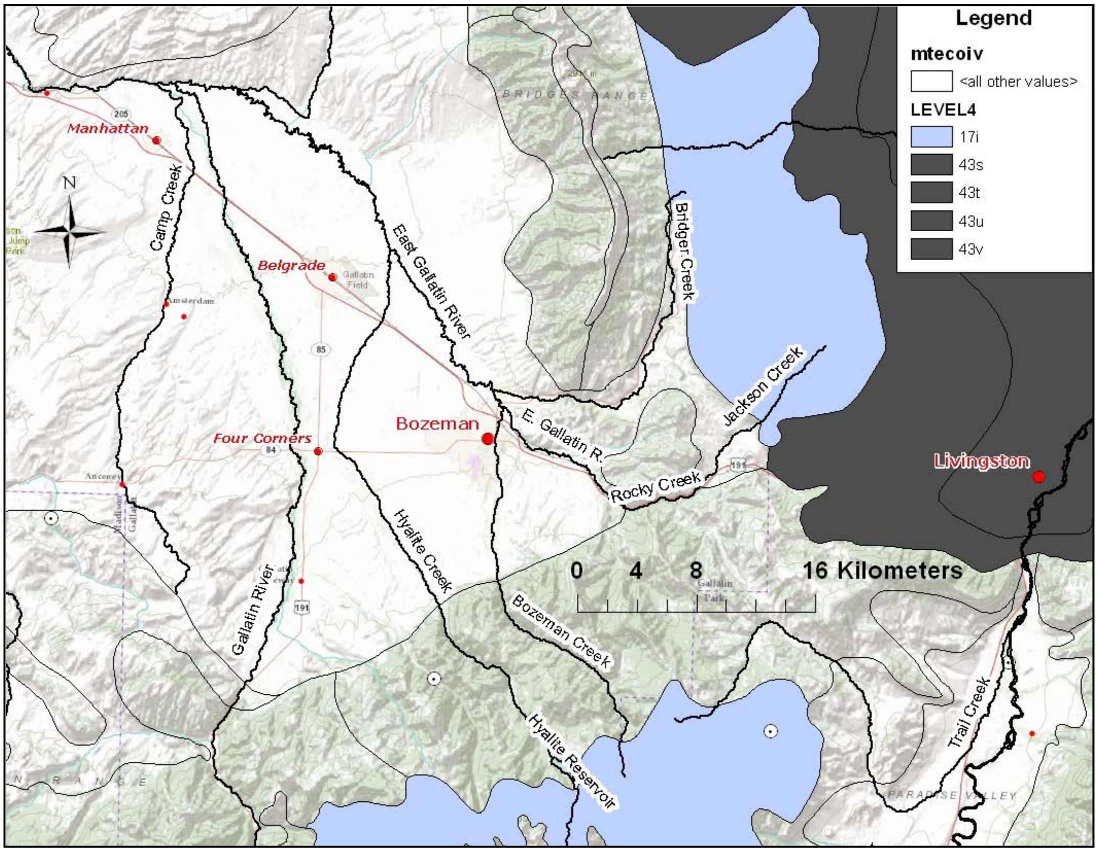

Figure 4-3. Map showing the East Gallatin River watershed, including Bozeman Creek and Hyalite Creek.

Blue shaded areas denote the level IV ecoregion 17i (Absaroka-Gallatin Volcanic Mountains).

Recommended Numeric Criteria

Table 4-2. Recommended Criteria for Reaches of Bozeman Creek, Hyalite Creek, and the East Gallatin River.

Stream Name	Reach Boundaries	TP Criterion (µg/L)	TN Criterion (µg/L)	TN:TP Ratio	Benthic Algal Biomass Criterion
Bozeman Creek	Headwaters to Forest Service Boundary (45.5833, -111.0184)	105	250	2:1	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Bozeman Creek	Forest Service Boundary (45.5833, -111.0184) to mouth at East Gallatin River	76	270	4:1	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Hyalite Creek	Headwaters to Forest Service Boundary (45.5833,-111.0835)	105	250	2:1	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
Hyalite Creek	Forest Service Boundary (45.5833,-111.0835) to mouth at East Gallatin River	90	260	3:1	125 mg Chl*a*/m^2 and 35 g AFDM/m^2

Table 4-2. Recommended Criteria for Reaches of Bozeman Creek, Hyalite Creek, and the East Gallatin River.

Stream Name	Reach Boundaries	TP Criterion (µg/L)	TN Criterion (µg/L)	TN:TP Ratio	Benthic Algal Biomass Criterion
East Gallatin River	Reach of East Gallatin River between Bozeman Creek and Bridger Creek confluences	50	290	6:1	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
East Gallatin River	Reach of East Gallatin River between Bridger Creek and Hyalite Creek confluences	40	300	8:1	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
East Gallatin River	Reach of East Gallatin River between Hyalite Creek and Smith Creek confluences	60	290	5:1	125 mg Chl*a*/m^2 and 35 g AFDM/m^2
East Gallatin River	Reach of East Gallatin River from Smith Creek confluence to the mouth (Gallatin River)	40	300	8:1	125 mg Chl*a*/m^2 and 35 g AFDM/m^2

Discussion of the Nutrient Criteria for Bozeman Creek, Hyalite Creek, and the East Gallatin River

In **Section 3.1.1**, we recommended TP and TN criteria specific to the Absaroka-Gallatin-Volcanic Mountains (17i), a level IV ecoregion with naturally elevated phosphorus concentrations. 'Elevated' means that the phosphorus levels in the ecoregion's reference streams were higher than the Middle Rockies (17) as a whole, and are naturally higher than concentrations that dose-response studies (phosphorus as cause, impact to stream beneficial use as effect) applicable to western Montana indicate are protective of beneficial uses.

The Hyalite Creek and Bozeman Creek watersheds contain parts of 17i, have documented elevated TP concentrations in surface water, and mapped Phosphoria formations within their boundaries (United States Geological Survey, 1951). Hyalite Creek and Bozeman Creek are in adjoining drainages and flow northward before joining the East Gallatin River (**Figure 4-3**). Bozeman Creek flows into the East Gallatin River at Bozeman, MT and Hyalite Creek joins the East Gallatin River northeast of Belgrade, MT. The headwaters of Jackson and Bridger creeks also fall within 17i, but that particular area does not have identified geologic sources of phosphorus or water quality data that suggest elevated phosphorus concentrations in surface water, and are not included in this discussion.

Reach-specific Methods

Nutrient data at the 75th percentile of reference for the Absaroka-Gallatin-Volcanic Mountains (17i) and the Middle Rockies (17) were used to determine the potential natural background of streams that flow through both ecoregions, and for waterbodies that receive drainage from both ecoregions (**Table 4-3**). Relative flow contributions were calculated from available discharge data from the USGS and from flow sampling projects conducted by the Department and its contractors. These flow estimates were used to determine the relative contribution from each ecoregional zone and, in turn, determine the potential natural background nutrient concentrations of each stream or stream segment using the following equation:

$$NB_{\mathbf{NEW}} = \frac{(NB_1 * Q_1) + (NB_2 * Q_2)}{Q_1 + Q_2}$$

Where NB_1 is the nutrient concentration (either N or P; µg/L) at the 75th percentile of the reference sites for ecoregion 17i, NB_2 is the nutrient concentration (either N or P; µg/L) at the 75th percentile of the reference sites for ecoregion 17 (Middle Rockies), Q1 and Q2 are the average summer flows (L/sec) that can be allocated to each ecoregional zone, and $NB_{\mathbf{NEW}}$ is the calculated natural-background nutrient concentration (µg/L) for the stream after having accounted for the mixing of the two water sources.

If the calculated natural background concentration (NB_{new}) in a given stream was equal to or greater than the recommended N or P criteria for the ecoregion in which the stream resides, a site-specific analysis was used to calculate the new criterion based on the estimated flow contributions from the different ecoregions. The new criterion was then derived using the mixing equation given above and using the draft ecoregional criteria (**Table 4-3**).

Table 4-3. Ecoregion-specific Reference Conditions and Numeric Nutrient Criteria for TN and TP. Data are from the all-observations dataset after applying the Brillouin Evenness Index.

	75th percentile - Reference Condition		Draft Numeric Nutrient Criteria	
	TN	**TP**	**TN**	**TP**
Level III Middle Rockies	141	20	300	30
Level IV Absaroka-Gallatin-Volcanics	100	105	250	105

All values are in µg/L

For example, in Bozeman Creek discharge records established that 63.4% of the flow at the mouth (1313.86 L/sec) originates upstream of the forest boundary (green area **in Figure 4-3**) where ecoregion 17i's TP concentrations are above the natural background for the Middle Rockies ecoregion. The balance of flow (36.6%; 481 L/sec) originates from below the forest boundary and is therefore associated with the Middle Rockies. Natural background (NB) for TP was calculated as:

$$([105\ \mu g\ TP/L * 833\ L/sec] + [20\ \mu g\ TP/L * 481\ L/sec]) \div (833 + 481\ L/sec) = 74\ \mu g\ TP/L$$

Because 74 µg TP/L exceeds the Middle Rockies criterion of 30 µg TP/L, a reach-specific criterion was then calculated for TP using the ecoregional numeric criteria:

$$([105\ \mu g\ TP/L * 833\ L/sec] + [30\ \mu g\ TP/L * 481\ L/sec]) \div (833 + 481\ L/sec) = 76\ \mu g\ TP/L$$

Criteria for reaches further downstream are then a function of concentrations and the proportion of flow coming from the upstream reach and the concentrations and flow from the tributary that demarcates the upper bound of the new reach. As before, it is a two-step process where estimated natural background is first calculated, and if the result exceeds the local ecoregional criterion, a reach-specific criterion is determined as a function of the criteria already derived for the two upstream waterbodies. This process can be carried downstream as far as needed. For example, for the reach "East Gallatin River between Hyalite Cr and Smith Cr" the TP criterion was calculated as follows:

$$(80\ \mu g\ TP/L * 0.325) + (30\ \mu g\ TP/L * 0.675) = 46\ \mu g/\ TP/L$$

Where 80 µg/TP and 0.325 are the calculated natural background for lower Hyalite Creek and its proportional contribution to flow in the new reach, respectively, and 30 µg TP/L is the calculated natural background concentration for the East Gallatin River just upstream of Hyalite Creek and 0.675 is its proportion of flow contribution to the new reach (**Table 4-4**). Since the calculated value of 46 µg TP/L exceeds the Middle Rockies regional criterion of 30 µg TP/L, the reach-specific criterion is then calculated:

(90 µg TP/L * 0.325) + (40 µg TP/L * 0.675) = 56.3 µg/ TP/L (rounds to 60 µg TP/L).

Results are shown below in **Table 4-4** for a subset of stream reaches in the area.

Table 4-4. Total Phosphorus Natural Background and Derived Nutrient Criteria for Example Stream and River Reaches in the East Gallatin River Watershed.

	Bozeman Creek (Forest Service boundary to mouth)	**East Gallatin R. between Bozeman and Bridger Creeks**	**East Gallatin R. between Bridger and Hyalite Creeks**	**Hyalite Creek (Forest Service boundary to mouth)**	**East Gallatin R. between Hyalite Cr and Smith Cr**
Natural Background	74	40	30	80	50
Reach Criterion	76	50	40	90	60

All values are in µg/L

Total phosphorus concentrations are directly affected by natural sources from ecoregion 17i in the Hyalite and Bozeman creek drainages (**Table 4-3**). Natural background for TP is at or above the numeric standard for the Middle Rockies ecoregion in every reach downstream of the phosphorus source area.

Data collected in 2008 and 2009 below the Bridger Creek confluence with the East Gallatin River but above the City of Bozeman WWTP discharge (*n*=5) had a mean of 22 µg TP/L with a maximum of 26 µg TP/L. These data generally support the calculation in **Table 4-4** where it was estimated that natural background for the reach in question would not be above 30 µg TP/L.

For waterbodies receiving significant flows from ecoregions with natural sources of phosphorus, adjusted downstream criteria for TN may be slightly lower based on the same equations and process described above (**Table 4-5**).

Table 4-5. Total Nitrogen Natural Background and Derived Criteria for Example Stream and River Reaches in the East Gallatin River Watershed.

Total Nitrogen	**Bozeman Creek (Forest Service boundary to mouth)**	**East Gallatin R. between Bozeman and Bridger Creeks**	**East Gallatin R. between Bridger and Hyalite Creeks**	**Hyalite Creek (Forest Service boundary to mouth)**	**East Gallatin R. between Hyalite Cr and Smith Cr**
Natural Background	120	130	140	110	130
Reach Criterion	270	290	300	260	290

All values are in µg/L

Note that for Bozeman and Hyalite creeks, the criteria applicable to ecoregion 17i (105 µg TP/L and 250 µg TN/L) apply to those streams from their respective headwaters down to the Forest Service boundary.

Conclusion

Equations relating benthic algal Chl*a* to total nutrients (Dodds et al., 1997, Equation 17; Dodds et al., 2006; Equation 19) were used to calculate the benthic Chl*a* biomass that would occur at the criteria levels shown for the stream and river reaches shown in **Table 4-2**. In all cases, benthic algae were maintained at ≤ 125 mg Chl*a*/m^2, therefore that value (and the accompanying AFDM value) is an appropriate and realistic level for these stream segments. The nutrient criteria are adequate to protect the coldwater fisheries use by assuring that dissolved oxygen levels always remains above standards at all times.

5.0 ACKNOWLEDGEMENTS

Many people contributed to the work that has led to this document. We thank Rosie Sada de Suplee, through whose work the Reference Stream Project was launched in 2000 and through whose efforts the identification and sampling of reference streams continues uninterrupted to this day. We thank the many Department of Environmental Quality employees who collected data at reference sites over the years. In particular, we thank Al Nixon who did an outstanding job of identifying and sampling reference sites particularly in the transitional zones of the Rocky Mountain Front. We thank the many field crew members from the University of Montana who collected data at stream reference sites around the state. The Montana Nutrient Work Group provided valuable feedback on the methods used to derive the numeric nutrient criteria. Finally, we would like to express our thanks to the many landowners around the state who provided us access to streams that ran through their lands.

6.0 REFERENCES

Borchardt, M. A. 1996. "Nutrients," in *Algal Ecology: Freshwater Benthic Ecosystems*, Stevenson, R. J., Bothwell, M. L., and Lowe, R. E., (San Diego, CA: Academic Press): 184-227.

Bothwell, M. L., C. Kilroy, B. W. Taylor, E. T. Ellison, D. A. James, C. A. Gillis, K. D. Bladon, and U. Silins. 2012. Iron Is Not Responsible for Didymosphenia Geminata Bloom Formation in Phosphorus-Poor Rivers. *Canadian Journal of Fisheries and Aquatic Sciences.* 69: 1723-1727.

Bothwell, M. L. 1989. Phosphorus-Limited Growth Dynamics of Lotic Periphytic Diatom Communities: Areal Biomass and Cellular Growth Rate Responses. *Canadian Journal of Fisheries and Aquatic Sciences.* 46(8): 1293-1301.

Bow River Basin Council. 2008. Bow Basin Watershed: Water Quality Objectives and Indicators. Bow Basin Watershed Management Plan Technical Committee.

Bowman, M. F., P. A. Chambers, and D. W. Schindler. 2007. Constraints on Benthic Algal Response to Nutrient Addition in Oligotrophic Mountain Rivers. *River Research and Applications.* 23(8): 858-876.

Bramblett, Robert G., Thomas R. Johnson, Alexander V. Zale, and Daniel Heggem. 2005. Development and Evaluation of a Fish Assemblage Index of Biotic Integrity for Northwestern Great Plains Streams. *Transactions of the American Fisheries Society.* 134(3): 624-640.

Brown, Claudeous J. D. 1971. Fishes of Montana, Bozeman, MT: Big Sky Books.

Caraco, N. F., J. J. Cole, and G. E. Likens. 1989. Evidence for Sulphate-Controlled Phosphorus Release From Sediments of Aquatic Systems. *Nature.* 341: 316-318.

Chambers, P. A., D. J. McGoldbrick, R. B. Brau, C. Vis, J. M. Culp, and G. A. Benoy. 2011. Development of Environmental Thresholds for Nitrogen and Phosphorus in Streams. *Journal of Environmental Quality.*

Chambers, P. A., E. E. Prepas, M. L. Bothwell, and Hal R. Hamilton. 1989. Roots Versus Shoots in Nutrient Uptake on Aquatic Macrophytes in Flowing Waters. *Canadian Journal of Fisheries and Aquatic Sciences.* 46: 435-439.

Conley, D. J., H. W. Paerl, R. W. Howarth, D. F. Boesch, S. P. Seitzinger, K. E. Havens, C. Lancelot, and G. E. Likens. 2009. Controlling Eutrophication: Nitrogen and Phosphorus. *Science (Washington).* 323: 1014-1725.

Dodd, W. K. and E. B. Welch. 2000. Establishing Nutrient Criteria in Streams. *Journal of the North American Benthological Society.* 19: 186-196.

Dodds, Walter K. 2003. Misuse of Inorganic N and Soluble Reactive P Concentrations to Indicate Nutrient Status of Surface Waters. *Journal of the North American Benthological Society.* 22(2): 171-181.

Dodds, Walter K. and R. M. Oakes. 2004. A Technique for Establishing Reference Nutrient Concentrations Across Watersheds Affected by Humans. *Limnology and Oceanography.* Methods 2: 333-341.

Dodds, Walter K., V. H. Smith, and K. Lohman. 2002. Nitrogen and Phosphorus Relationships to Benthic Algal Biomass in Temperate Streams. *Canadian Journal of Fisheries and Aquatic Sciences.* 59: 865-874.

-----. 2006. Erratum: Nitrogen and Phosphorus Relationships to Benthic Algal Biomass in Temperate Streams. *Canadian Journal of Fisheries and Aquatic Sciences.* 63: 1190-1191.

Dodds, Walter K., V. H. Smith, and Bruce Zander. 1997. Developing Nutrient Targets to Control Benthic Chlorophyll Levels in Streams: A Case Study of the Clark Fork River. *Water Resources.* Vol. 31(no. 7): 1738-1750.

Elser, J. J., M. E. S. Bracken, E. E. Cleland, D. S. Gruner, W. S. Harpole, H. Hillebrand, J. T. Ngai, E. W. Seabloom, J. B. Shurin, and J. E. Smith. 2007. Global Analysis of Nitrogen and Phosphorus Limitation of Primary Producers in Freshwater, Marine and Terrestrial Ecosystems. *Ecology Letters.* 10(12): 1135-1142.

Elser, J. J., E. R. Marzolf, and C. R. Goldman. 1990. Phosphorus and Nitrogen Limitation of Phytoplankton Growth in the Freshwaters of North America: a Review and Critique of Experimental Enrichments. *Canadian Journal of Fisheries and Aquatic Sciences.* 47(7): 1468-1477.

Flynn, K. and Michael W. Suplee. 2010. Defining Large Rivers in Montana Using a Wadeability Index. Helena, MT: Montana Department of Environmental Quality. http://deq.mt.gov/wqinfo/Standards/default.mcpx.

Flynn, Kyle and Michael W. Suplee. 2013. Using a Computer Water Quality Model to Derive Numeric Nutrient Criteria: Lower Yellowstone River - Final. Helena, MT: Montana Department of Environmental Quality. WQPBMSTECH-22.

Francoeur, S. N. 2001. Meta-Analysis of Lotic Nutrient Amendment Experiments: Detecting and Quantifying Subtle Responses. *Journal of the North American Benthological Society.* 20(3): 358-368.

Gibson, C. E. 1971. Nutrient Limitation. *Journal of the Water Pollution Control Federation.* 43: 2436-2440.

Gravelle, J. A., G. Ice, T. E. Link, and D. L. Cook. 2009a. Nutrient Concentration Dynamics in an Inland Pacific Northwest Watershed Before and After Timber Harvest. *Forest Ecology and Management.* 257: 1663-1675.

Gravelle, J. A., T. E. Link, J. R. Broglio, and J. H. Braatne. 2009b. Effects of Timber Harvest on Aquatic Macroinvertebrate Community Composition in a Northern Idaho Watershed. *Forest Science.* 55(4): 352-366.

Hasler, A. D. and W. G. Einsele. 1948. Fertilization for Increasing Productivity of Natural Inland Waters. In: Trans.North American Wildlife Conferece.   527-555.

Hecky, R. E. and P. Kilham. 1988. Nutrient Limitation of Phytoplankton in Freshwater and Marine Environments: A Review of Recent Evidence on the Effects of Enrichment. *Limnology and Oceanography.* 33: 796-822.

Heiskary, Steven, R. W. Bouchard, and H. Markus. 2010. Minnesota Nutrient Criteria Development for Rivers. Saint Paul, MN: Minnesota Pollution Control Agency. http://www.pca.state.mn.us/index.php/view-document.html?gid=14947.

Hillebrand, H. and U. Sommer. 1999. The Nutrient Stoichiometry of Benthic Microalgal Growth: Redfield Proportions Are Optimal. *Limnology and Oceanography.* 44: 440-446.

Holderman, C., G. Hoyle, R. Hardy, P. Anders, P. Ward, and H. Yassien. 2009. Libby Dam Hydro-Electric Project Mitigation: Efforts for Downstream Ecosystem Restoration. In: 33rd IAHR Congress: Water Engineering for a Sustainable Envrionment. 33rd IAHR Congress: Aug. 9, 2009; Vancouver, BC; 12986.

Hooker, H. D. 1917. Liebig's Law of the Minimum in Relation to General Biological Problems. *Science.* 46: 197-204.

Horner, R. R., E. B. Welch, and R. B. Veenstra. 1983. "Development of Nuisance Periphytic Algae in Laboratory Streams in Relation to Enrichment and Velocity," in *Periphyton of Freshwater Ecosytems*, Wetzel, R. G., (The Hague: Dr. W. Junk Publishers)

Hudon, C. and E. Bourget. 1981. Initial Colonization of Artificial Substrate: Community Development and Structure Studied by Scanning Electron Microscopy. *Canadian Journal of Fisheries and Aquatic Sciences.* 38(11): 1371-1384.

Hullar, M. A. and J. R. Vestal. 1989. The Effects of Nutrient Limitation and Stream Discharge on the Epilithic Microbial Community in an Oligotrophic Arctic Stream. *Hydrobiologia.* 172(1): 19-26.

Hurlbert, S. H. 1984. Pseudoreplication and the Design of Ecological Field Experiments. *Ecological Monographs.* 54: 187-211.

Jewell, W. J. 1971. Aquatic Weed Decay: Dissolved Oxygen Utilization and Nitrogen and Phosphorus Regeneration. *Journal (Water Pollution Control Federation).* 43: 1457-1467.

Johnston, N. T., C. J. Perrin, P. A. Slaney, and B. R. Ward. 1990. Increased Juvenile Salmonid Growth by Whole-River Fertilization. *Canadian Journal of Fisheries and Aquatic Sciences.* 47(5): 862-872.

Kilroy, C. 2011. Environmental Control of Stalk Length in the Bloom-Forming, Freshwater Benthic Diatom Didymosphenia Geminata (Bacillariophyceae).

Kilroy, C. and M. L. Bothwell. 2012. Didymosphenia Geminata Growth Rates and Bloom Formation in Relation to Ambient Dissolved Phosphorus Concentration. *Freshwater Biology.* 57(4): 641-653.

KOHLER, A. E., A. RUGENSKI, and D. TAKI. 2008. Stream Food Web Response to a Salmon Carcass Analogue Addition in Two Central Idaho, USA Streams. *Freshwater Biology.* 53(3): 446-460.

Lewis, William M. and Wayne A. Wurtsbaugh. 2008. Control of Lacustrine Phytoplankton by Nutrients: Erosion of the Phosphorus Paradigm. *International Review of Hydrobiology.* 93(4-5): 446-465.

Lewis, William M. J., Wayne A. Wurtsbaugh, and Hans W. Paerl. 2011. Rationale for Control of Anthropogenic Nitrogen and Phosphorus to Reduce Eutrophication of Inland Waters. *Environmental Science & Technology.* 45(24): 10300-10305.

Lovstad, O. 2008. A Phosphorus-Based Biological Classification System and Threshold Indicators. *Internationale Vereinigung Für Theoretische Und Angewandte Limnologie.* 30: 565-568.

Lyons, J. 1992. Using the Index of Biotic Integrity (IBI) to Measure Environmental Quality in Warmwater Streams of Wisconsin, General Technical Report NC-149 ed., St. Paul, MN: US Department of Agriculture, Forest Service, North Central Forest Experiment Station St. Paul, MN. Accessed 5/7/12.

Mebane, C. A., A. M. Ray, A. M. Marcarelli, and Flint Raben. 2009. Influence of Nutrients on Agricultural Stream Ecosystems: Integrating Biomonitoring and Experimental Information. In: 20th Annual Northwest Bioassessment Workgroup Meeting. Nov. 4, 2009; McCall, ID.

Montana Department Environmental Quality. 2009. Box Elder Creek Nutrient-Addition Project: 2009 Preliminary Data Collection Sampling and Analysis Plan. Helena, MT: Montana Department of Environmental Quality.

-----. 2011. Periphyton Standard Operating Procedure. WQPBWQM-010.

Montana Department of Environmental Quality. 2011. Sample Collection and Laboratory Analysis of Chlorophyll-*a* Standard Operation Procedure, Revision 5. Helena, MT: Montana Department of Environmental Quality. WQPBWQM-011.

Mortimer, C. H. 1941. The Exchange of Dissolved Substances Between Mud and Water in Lakes. *Journal of Ecology.* 29(3 and 4): 280-329.

Omernik, James M. 1987. Ecoregions of the Conterminous United States. *Annals of the Association of American Geographers.* 77: 118-125.

Paerl, H. W. 2009. Controlling Eutrophication Along the Freshwater-Marine Continuum: Dual Nutrient (N and P) Reductions Are Essential. *Estuaries and Coasts.* 32(4): 593-601.

Perrin, C. J., M. L. Bothwell, and P. A. Slaney. 1987. Experimental Enrichment of a Coastal Stream in British Columbia: Effects of Organic and Inorganic Additions on Autotrophic Periphyton Production. *Canadian Journal of Fisheries and Aquatic Sciences.* 44: 1247-1256.

Perrin, C. J. and J. S. Richardson. 1997. N and P Limitation of Benthos Abundance in the Nechako River, British Columbia. *Canadian Journal of Fisheries and Aquatic Sciences.* 54(11): 2574-2583.

Pielou, E. C. 1966. The Measurement of Diversity in Different Types of Biological Collections. *Journal of Theoretical Biology.* 13: 131-144.

Rier, S. T. and R. J. Stevenson. 2006. Response of Periphytic Algae and Gradients in Nitrogen and Phosphorus in Streamside Mesocosms. *Hydrobiologia.* 561: 131-147.

Scrimgeour, G. J. and P. A. Chambers. 2000. Cumulative Effects of Pulp Mill and Municipal Effluents on Epilithic Biomass and Nutrient Limitation in a Large Northern River Ecosystem. *Canadian Journal of Fisheries and Aquatic Sciences.* 57(7): 1342-1354.

Smith, Val H., G. D. Tilman, and J. C. Nekola. 1999. Eutrophication: Impacts of Excess Nutrient Inputs on Freshwater, Marine, and Terrestrial Ecosystems. *Environmental Pollution.* 100(1999): 176-196.

Sosiak, A. 2002. Long-Term Response of Periphyton and Macrophytes to Reduced Municipal Nutrient Loading to the Bow River (Alberta, Canada). *Canadian Journal of Fisheries and Aquatic Sciences.* 59: 987-1001.

Spaulding, Sarah and Leah Elwell. 2007. Increase in Nuisance Blooms and Geographic Expansion of the Freshwater Diatom Didymosphenia Geminata: Recommendations for Response. USGS Open File Report 2007-1425. http://prdp2fs.ess.usda.gov/Internet/FSE_DOCUMENTS/fsbdev3_015009.pdf. Accessed 5/7/2012.

Stevenson, R. J., Steven T. Rier, Catherine M. Riseng, Richard E. Schultz, and Michael J. Wiley. 2006. Comparing Effects of Nutrients on Algal Biomass in Streams in Two Regions With Different Disturbance Regimes and With Applications for Developing Nutrient Criteria. *Developments in Hydrobiology.* 185: 149-165.

Stockner, J. 2003. Nutrients in Salmonid Ecosystems: Sustaining Production and Biodiversity. 2001 Nutrient Conference, Restoring Nutrients of Salmonid Ecosytems. In: Stockner, J. G. (ed.). Proceedings of the 2001 Nutrient Conference, Restoring Nutrients of Salmonid Ecosytems. 2001 Nutrient Conference Restoring Nutrients of Salmonid Ecosytems: American Fisheries Society Symposium. Bethseda, MD: American Fisheries Society.

Stockner, J. G. and K. I. Ashley. 2003. Salmon Nutrients: Closing the Circle. In: Stockner, J. G. (ed.). Nutrients in Salmonid Ecosystems: Sustaining Production and Biodiversity.2001 Nutrient Conference, Restoring Nutrients of Salmonid Ecosytems. 2001 Nutrient Conference Restoring Nutrients of Salmonid Ecosytems: American Fisheries Society Symposium. Bethseda, MD: American Fisheries Society; 3-15.

Stockner, J. G. and K. R. S. Shortreed. 1978. Enhancement of Autotrophic Production by Nutrient Addition in a Coastal Rainforest Stream on Vancouver Island. *Journal of the Fisheries Board of Canada.* 35(1): 28-34.

Strahler, A. N. 1964. "Quantitative Geomorphology of Drainage Basins and Channel Networks," in *Handbook of Applied Hydrology*, Chow, V. T., (New York: McGraw-Hill): 439-476.

Streeter, H. W. and E. B. Phelps. 1925. A Study of the Pollution and Natural Purfication of the Ohio River, III. Factors Concerning the Phenomena of Oxidation and Reaeration. *Public Health Bulletin.* 146(February 1925)

Sundareshwar, P. V., S. Upadhayay, M. Abessa, S. Honomichl, B. Berdanier, S. A. Spaulding, C. Sandvik, and A. Trennepohl. 2011. Didymosphenia Geminata: Algal Blooms in Oligotrophic Streams and Rivers. *Geophysical Research Letters.* 38

Suplee, Michael W. 2004. Wadeable Streams of Montana's Hi-Line Region : An Analysis of Their Nature and Condition With an Emphasis on Factors Affecting Aquatic Plant Communities and Recommendations to Prevent Nuisance Algae Conditions. Helena, MT: Montana Department of Environmental Quality, Water Quality Standards Section.

Suplee, Michael W. and J. B. Cotner. 2002. An Evaluation of the Importance of Sulfate Reduction and Temperature to P Fluxes From Aerobic-Surfaced, Lacustrine Sediments. *Biogeochemistry.* 61: 199-228.

Suplee, Michael W. and Rosie Sada de Suplee. 2011. Assessment Methodology for Determining Wadeable Stream Impairment Due to Excess Nitrogen and Phosphorus Levels. Helena, MT: Montana Department of Environmental Quality Water Quality Planning Bureau. WQPMASTR-01.

Suplee, Michael W., Rosie Sada de Suplee, David L. Feldman, and Tina Laidlaw. 2005. Identification and Assessment of Montana Reference Streams: A Follow-Up and Expansion of the 1992 Benchmark Biology Study. Helena, MT: Montana Department of Environmental Quality.

Suplee, Michael W., Arun Varghese, and Joshua Cleland. 2007. Developing Nutrient Criteria for Streams: An Evaluation of the Frequency Distribution Method. *Journal of the American Water Resources Association.* 43(2): 456-472.

Suplee, Michael W., V. Watson, M. Teply, and H. McKee. 2009. How Green Is Too Green? Public Opinion of What Constitutes Undesirable Algae Levels in Streams. *Journal of the American Water Resources Association.* 45: 123-140.

Suplee, Michael W., V. Watson, A. Varghese, and Joshua Cleland. 2008. Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers. Helena, MT: MT DEQ Water Quality Planning Bureau.

Suplee, Michael W., Vicki Watson, Walter K. Dodds, and Chris Shirley. 2012. Response of Algal Biomass to Large Scale Nutrient Controls in the Clark Fork River, Montana, U.S.A. *Journal of American Water Resources Association.* 48: 1008-1021.

Tank, J. L. and W. K. Dodds. 2003. Response of Heterotrophic and Autotrophic Biofilms to Nutrients in Ten Streams. *Freshwater Biology.* 48: 1031-1049.

Teply, M. 2010. Interpretation of Periphyton Samples From Montana Streams. Cramer Fish Sciences. Helena, MT: Montana Department of Environmental Quality.

Teply, M. and L. Bahls. 2007. Statistical Evaluation of Periphyton Samples From Montana Reference Streams. Larix Systems Inc. and Hannaea. Helena, MT: Montana Department of Environmental Quality.

United States Geological Survey. 1951. Stratigraphic Sections of the Phosphoria Formation in Montana, 1951. *Geological Servey Circular.* 326

Varghese, A. and Joshua Cleland. 2008. Updated Statistical Analyses of Water Quality Data, Compliance Tools, and Changepoint Assessment for Montana Rivers and Streams. Fairfax, VA.

Varghese, Arun and Joshua Cleland. 2005. Seasonally Stratified Water Quality Analysis for Montana Rivers and Streams: Final Report. Fairfax, VA: ICF Consulting.

-----. 2009. Creation of an Integrated Lake Water Quality Dataset, Updating of the Rivers & Streams Dataset, and Statistical Analysis of the Updated Rivers & Streams Data. ICF International. DEQ Contract #205031.

Wang, L., D. M. Robertson, and P. J. Garrison. 2007. Linkages Between Nutrients and Assemblages of Macroinvertebrates and Fish in Wadeable Streams: Implication to Nutrient Criteria Development. *Environmental Management.* 39: 194-212.

Welch, E. B., R. R. Horner, and C. R. Patmont. 1989. Prediction of Nuisance Periphytic Biomass: a Management Approach. *Water Research.* 23: 401-405.

Welcomme, R. L. 1985. River Fisheries. Rome, Italy: United Nations Food and Agricultural Organization. FAO Fisheries Technical Papaer No. 262. http://www.fao.org/DOCREP/003/T0537E/T0537E03.htm#ch3. Accessed 5/7/2012.

Whitton, B. A., N. T. W. Ellwood, and B. Kawecka. 2009. Biology of the Freshwater Diatom Didymosphenia: A Review. *Hydrobiologia.* 630(1): 1-37.

Woods, A. J., James M. Omernik, J. A. Nesser, J. Shelden, J. A. Comstock, and S. J. Azevedo. 2002. Ecoregions of Montana, 2nd ed., Reston, VA: United States Geographical Survey.

Zarr, J. H. 1999. Biostatistical Analysis. *Prentice Hall New Jersey.* 4

APPENDIX A – PEER-REVIEW COMMENTS AND RESPONSES

Montana Department of
ENVIRONMENTAL QUALITY

MEMO

1520 E 6th Avenue
PO Box 200901
Helena MT 59620-0901
(406) 444-6697 • FAX: (406) 444-3836

To: Tina Laidlaw, U.S Environmental Protection Agency

CC: Eric Urban, Head, Water Quality Standards Section

From: Michael Suplee, Ph.D., Water Quality Standards Section; Vicki Watson, Ph.D., University of Montana

Date: 5/9/2013

RE: MT DEQ's response to peer-review comments on "Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers: Addendum 1"

Reviews were received by three anonymous peer reviewers on the document referenced above in August, 2012. The reviewers were selected by the U.S. EPA in conjunction with the Nutrient Scientific Technical Exchange Partnership & Support (NSTEPS) service. One of the services NSTEPS provides is review of state-developed numeric nutrient criteria.

Section 1.0 below addresses comments that were common to two or all three reviewers; MT DEQ's response is provided in each case. **Section 2.0** lists salient comments from individual reviewers. **Section 3.0** summarizes changes to the "Addendum 1" document[1] that will be made as a result of reviewers' comments. Some comments were minor or editorial in nature and these have simply been addressed during the finalization of the document.

1.0 Comments from Peer Reviewers Addressing EPA's Six Core Questions

EPA posed six questions to the reviewers. The first queried their overall impression of the approach MT DEQ took to derive the numeric nutrient criteria. There was universal agreement among the three reviewers that the approach taken was thorough, scientifically sound, and an effective use of available and relevant information. There were, of course, concerns and recommendations as well. The five remaining EPA questions are addressed in each of the sections below (**Section 1.1** to **Section 1.5**) and most of the reviewer's comments/concerns are covered in these sections. **Sections 1.6 and 1.7** address other issues raised by the peer reviewers.

---

[1] The draft document was called "Addendum 1" because we considered it an extension of methods and ideas put forth in Suplee et al. (2008). However, enough material has changed and the document is now sufficiently stand-alone that in final form it has been named "Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers: Update 1".

1.1 Concern the MT DEQ has not Provided Nutrient Criteria Recommendations for the Level IV Ecoregion "River Breaks"

Two reviewers were concerned that MT DEQ did not provide draft criteria for this ecoregion. They wanted to see more reference sites, recommended that MT DEQ discuss soluble nutrients from the ecoregion's reference sites, and discuss the potential impact on downstream uses if no criteria were adopted in this area. A third reviewer was apparently very familiar with western plains environments, and understood our reasoning, but was still concerned about downstream use impacts.

RESPONSE: The basic tenant of MT DEQ's approach is to apply appropriate stressor-response studies to a region and then compare the harm-to-use thresholds derived from the studies to the reference distribution (Suplee et al., 2007). Although there are just eight reference sites in the River Breaks, it is not so small a dataset as to preclude reasonable comparisons to dose-response studies. Compiling the reference data by site medians (see discussion on this topic in **Section 1.6** below) did not substantially alter the plains region dose-response-to-reference matches (equal to the 48th percentile for the median dataset and the 53rd percentile for the all-observations dataset, for TP; 60th and the 53rd,respectively, for TN). These data show, regardless of how the reference data are summarized, that harm-to-use concentrations applicable to the plains align with nutrient concentrations which are about average in the River Breaks' reference streams (i.e., River Breaks streams are naturally eutrophied or may have already responded to global increases in nitrogen loading [Vitousek et al., 1997]).

Per reviewers' recommendations, we have included soluble nutrients in the final report for the River Breaks. In relation to stream algal growth, it results that soluble nutrients in the River Breaks reference streams are already saturated, or are nearly so. For NO_2+NO_3, the 75th percentile of River Breaks reference streams was 241 or 631 µg N/L (median or all-observations datasets, respectively). Rier and Stevenson (2006) show there is little peak algal biomass increase above 308 µg soluble nitrogen per liter (and peak biomass may actually be saturated closer to 250 µg DIN/L). As such, River Breaks reference sites are often saturated with soluble nitrogen, which is the nutrient most likely to be added to these streams if future development were to occur. Soluble P concentrations are also high. At the 75th percentile of reference SRP is 18 or 20 µg P/L (median vs. all-observations datasets, respectively) and therefore these low-gradient reference streams are often P saturated for peak algal biomass, or nearly so (Horner et al., 1983; Bothwell, 1989).

The absence of numeric nutrient standards in the River Breaks does not mean there will be no nutrient controls whatsoever applied to new permitted sources. Aquatic life ammonia standards still apply year-round. Median pH in the River Breaks reference streams is 8.8, and with typical summer temperatures of 20 to 25°C, the ammonia criterion would be about 360 µg NH_{3+4}-N/L (DEQ-7, 2012) and would provide protection from the toxic effects of ammonia on early fish life stages. If all this ammonia were oxidized to nitrate the resulting nitrate concentration would be well within the nitrate range observed in the River Breaks reference sites. The human health standards of 1.0 mg NO_2-N/L and 10 mg NO_3-N/L would apply year round as well. Thus, Montana's existing water quality standards would preclude the River Breaks from becoming an 'industrial dumping ground', a concern expressed by one reviewer.

Downstream uses will be addressed in permitting situations via application of nondegradation. The River Breaks ecoregion basically drains directly into the mainstem Missouri and Yellowstone rivers. In spite of the elevated nitrates and total N and P coming from the River Breaks, MT DEQ has not observed nutrient problems in the lower Yellowstone River, i.e., algal levels at unacceptable levels or DO and pH that violate state quality standards. Summertime concentrations in the Yellowstone River near Glendive (in the heart of River Breaks country) during low-flow years average 490 µg TN/L and 55 µg TP/L, and are well below our recommended numeric nutrient criteria for the lower Yellowstone River during low flow(815 µg TN/L and 95 µg TP/L; Flynn and Suplee, 2013). In establishing any permit which would allow an N or P discharge that is likely to reach the Yellowstone or Missouri River, nondegradation would be considered.

We conclude that there is no scientifically-defensible way to derive numeric nutrient criteria for the control of eutrophication for streams of the River Breaks. The streams are highly turbid, flashy, have low levels of benthic algae and macrophytes, and have soluble nutrient concentrations at levels that saturate algal growth much of the time. Other MT DEQ programs will address impacts to downstream uses. We will not be recommending nutrient criteria for these streams in the final report.

1.2 Peer Reviewers' Views Concerning the Allowable 20% Exceedence Rate Associated with the Criteria (Pertains to Assessment Methodology[2])

RESPONSE: This topic closely ties to the topic in **Section 1.5** below, and is addressed there.

1.3 Peer Reviewers' Comments on MT DEQ's Use of Benthic Chlorophyll *a*, How the Chlorophyll *a* Threshold (125 mg Chl*a*/m^2) was Derived, and Thoughts on Other Biological Measurements Used to Support Eutrophication Assessment

Two reviewers found the use of benthic chlorophyll *a* to be an excellent tool for assessing eutrophication, while the third did not like it. Derivation of the chlorophyll *a* thresholds were considered appropriate although two reviewers felt that the threshold of 125 mg Chl*a*/m^2 may be too close to the harm-to-use threshold. (The third reviewer found it acceptable.) One reviewer notes that macroinvertebrates are a poor indicator of eutrophication.

RESPONSE: MT DEQ has had good success with measuring benthic chlorophyll *a* and does not believe the concerns of one reviewer (it's too variable, affected by grazers) apply to the physiographic regions where it is used. Note also that MT DEQ collects benthic ash free dry mass (as g/m^2), which can provide good indication of heavy benthic algal growth even if chlorophyll *a* levels have declined due to senescence. As pointed out by one reviewer, MT DEQ has a long tradition of measuring benthic algae density and diatom taxa and both of these are main features of the nutrient assessment method. We agree that macroinvertebrates are not an ideal tool for pinpointing eutrophication problems, which is why they are used secondarily, i.e., only after the better tools (benthic chlorophyll *a*, diatom metrics) have already been played out. At a recent conference of academic experts on stream ecology (April 16-18, 2013, Washington, D.C.), which one of the authors was fortunate enough to attend, there was wide agreement that macroinvertebrates have generally poor predictive power for eutrophication assessment.

MT DEQ had extensive internal discussion about where to set the benthic algae density after the results from the dosing study (Appendix B, Suplee and Sada de Suplee, 2011) showed that average levels of 127

---

[2] MT DEQ's assessment methodology for assessing eutrophication in wadeable streams (Suplee and Sada de Suplee, 2011) was completely revised in 2009-2010, went through public comment (including EPA review), and was finalized prior to the time that it was provided to the peer reviewers here.

mg Chl*a*/m^2 could result in seasonal DO problems. To inform that discussion, a mechanistic model was built to simulate the DO impact observed in the dosing study and the model showed that higher gradient streams in western Montana would not develop low DO due to their reaeration; lower gradient streams, however, would be impacted. In streams with good re-aeration, therefore, harm-to-use would not occur until 150 mg Chla/m^2 (the recreational threshold). To avoid creating an overly-complex application of the algae threshold, involving not only ecoregions but different beneficial uses and different benthic algae levels for different stream gradients, it was decided that one algae threshold would be established (125 mg Chl*a*/m^2) that should be largely protective of both aquatic life and recreation. Monitoring staff with experience using the thresholds understood the rationale and indicated that they were comfortable with it because, in most cases, streams' algae densities are well below or well above the thresholds, precluding borderline decisions. MT DEQ is measuring diel DO concentrations much more frequently now and will continue to evaluate the 125 mg Chl*a*/m^2 threshold; it can be readjusted if needed in the future.

1.4 Peer Reviewers' Assessment of MT DEQ's Reach-specific Nutrient-criteria Derivation Method

All three reviewers supported the approach taken.

RESPONSE: We are delighted that all three peer reviewers were very supportive of the approach that was taken.

1.5 Peer Reviewers' Views Concerning the use of the Binomial Test, its 20% Allowable Exceedence Rate, and the Student's T-test (Pertains to Assessment Methodology)

The reviewers' main questions and thoughts/observations pertaining to these subjects are summarized as follows. Questions: (a) By using an allowable exceedence rate of 20% and an effect size of 15%, is the exact binomial testing whether 35% of observations must exceed the criterion to be considered non-compliant? (b) Is it appropriate to use an effect size in the T-test? Thoughts/observations: (a)The T-test is a parametric test with assumptions of a normal distribution which are not the norm for the datasets being evaluated, and so it will be less likely to detect a difference in the mean of a nutrient dataset relative to the criterion, and (b) for the T-test to establish non-compliance, the average concentration of a test stream would need to be substantially above the 75th percentile of reference (the reviewer's presumed level at which protection of uses is assured) and this is under protective.

In addition, there was general confusion among reviewers on how MT DEQ has defined an observation when assessing nutrients, how reaches are delineated, and how statistical tests and biological information all fit together in the final assessment.

RESPONSE: MT DEQ's statistical assessment of nutrient concentrations in a stream segment can be reduced to two simple ideas: (1) a test (exact binomial) to determine the proportion of samples that exceed the criterion, and (2) a test (one sample Student's T-test for the mean) to help to identify when the average nutrient concentration has been pulled above the criterion, which may result because most samples are above the criterion or because just a few high outliers are. Each test is discussed below, followed by an overall conclusion.

Binomial Test. Excellent empirical data were available to MT DEQ to derive the allowable exceedence rate used in the test (more on this in a moment). Besides the all-important exceedence rate, MT DEQ had to give consideration to other factors for statistical testing including the realistic number of independent samples that could be collected in a stream reach (restrained by cost/time), and the desire to balance type I and II error rates, i.e., give roughly equal weight to the importance of error of over-regulation vs. failure to protect the environment (Mapstone, 1995). Realistic sample sizes were about 10 to 15, and as such it was impossible to have alpha and beta error both around 0.05 (95% confidence) because sample size would then need to be around 75. So MT DEQ opted for less confidence, i.e., alpha and beta error rates both ≈0.25 (75% confidence).

Allowable exceedence rate (number of samples allowed above the criterion while assuring the river supports beneficial uses) was empirically derived from long-term work on the Clark Fork River—a river where adopted nutrient standards are virtually identical to those proposed for western Montana streams. **The Clark Fork River analysis shows that a defensible criteria exceedence rate could range from 5-31%.** Twenty percent was identified as the most reasonable value. To date, MT DEQ has not found or been made aware of another dataset by which an allowable exceedence rate for numeric nutrient criteria could be determined. Because of this, MT DEQ will continue to use the 20% exceedence rate.

MT DEQ uses a 15% effect size. By establishing 15% effect size, MT DEQ is saying that this is the range of true exceedence rates where the consequence of decision errors is relatively minor. As a point of comparison, if there were a pollutant for which the allowable exceedence rate is set at 10% and it is known that virtually no impact will occur at 9% exceedence, but terrible impacts occur at 11% exceedence, then the effect size would have to be set very close to zero, because the consequence of decision error is huge. And as a result, very large numbers of samples may need to be collected to discern with accuracy that fine a cut on the exceedence rate. But for nutrients, the state-of-the science is still limited and what we do know tells us there is a fairly wide range (5-31%) where decision error impacts are minor; MT DEQ addressed this by selecting a somewhat wide (15%) effect size.

In the binomial—with 20% allowable exceedence rate and 15% effect size—MT DEQ is establishing that streams with <5% exceedence will always PASS (be found compliant with) the binomial test, and streams with >35% exceedence will always FAIL (be found non-compliant with) the binomial. (The reviewer is correct that the 20% exceedence rate and 15% effect size are additive.) Streams falling in between will sometimes PASS, sometimes FAIL (depends on **n**). It could be reasonably argued that 35% exceedence is too high, but sample-size reality then enters the picture: if we lower the effect size to 10%, i.e. streams with 30% exceedence rate will always FAIL the binomial, we would have to collect 25 samples to roughly balance alpha and beta error; too many samples to institute for routine stream nutrient monitoring. Other combinations of exceedance rates and effect sizes within defensible ranges (and again balancing alpha, beta error) also led to **n**'s in the low 20s or higher. In the end, MT DEQ settled on the exceedence rate and effect size we are currently using. However, note that in borderline situations (i.e., the assessment decision is not clear) MT DEQ will collect more data, and may very well end up with sample sizes closer to 20.

T-test. Per the reviewer's question, no, effect size is not included in the T-test. MT DEQ believes that at this point the EPA-recommended T-test is satisfactory for its purpose within the assessment methodology. It is robust against moderate deviations from normality (and many of the small datasets that are considered are essentially normally distributed). The reviewer is correct that the T-test loses power when datasets are highly skewed (and some of the datasets are skewed). But in actual cases where there are a few very large outliers among the 12 or so samples (this is a common scenario), the T-test still FAILS (indicates non-compliance, as we would want it to) even if the exact test statistics (p value, etc.) may not be particularly accurate. Staff who routinely carry out eutrophication assessments have expressed that the T-test results are largely in alignment with the totality of information provided by the binomial and biological measurements.

Regarding the idea that the average concentration in a test site would have to be much greater than the 75th of reference in order to FAIL the T-test, two points can be made. (1) The same reviewer stated that nutrient criteria are best if based upon dose-response studies. MT DEQ has found that dose-response studies often show concentrations >75th of reference are protective of legally-defined beneficial uses (Suplee et al., 2007; Suplee et al., 2008). Thus, PASSING the T-test because the average concentration in a test site is >75th percentile of reference is not necessarily under protective. (2) MT DEQ uses a different test hypothesis depending on the stream's 303(d) listing history for nutrients. Already-listed streams have the null as "stream is impaired" and the alternative as "not impaired". Thus, MT DEQ has the most control on alpha error which is defined upfront in the test. This approach is more protective.

MT DEQ Procedures and Assumptions. Regarding clear explanations of MT DEQ procedures, MT DEQ laid out the entire assessment approach and its assumptions in Suplee and Sada de Suplee (2011), including a number of examples that can be followed (see Section 3.2.4 of that document). However it appears that the final element of the method, the data-review matrix contained in the Excel spreadsheet "NtrntAssessFramework.xlsx", may not have been seen by some reviewers. Lacking this final piece would have led to confusion for sure. In any case, MT DEQ believes that "Addendum 1" (now Update 1) is not the place to detail assessment methodologies that are well covered in other documents. Going forward, Update 1 will continue to focus on nutrient criteria and their derivation.

Conclusion. The binomial test and the T-test in MT DEQ's assessment methodology will continue to be used as configured. As noted by a reviewer, MT DEQ compensates for the higher-than-ideal FAIL threshold of 35% in the binomial test by establishing different null hypothesis depending on if the stream is (or is not) already listed on the 303(d) list, by including the T-test, and by lowering the chlorophyll *a* threshold to 125 mg Chla/m^2 (instead of 150 mg Chla/m^2). As noted by another reviewer, "Some of the quibbling on these values may never be resolved (including mine), and Montana needs to use best judgment supported by its analysis and other scientific results." We couldn't agree more.

1.6 Number of Reference Sites, Manner by which MT DEQ Characterizes the Reference Condition

One reviewer felt there were too few reference data. Reviewers felt that MT DEQ's novel use of the Brillouin Evenness Index should be (at a minimum) clearly spelled out, and include the equations. One reviewer felt that the Brillouin method was "interesting", but that it did not directly address the issue of temporal pseudoreplication which may arise in repeated measurements of nutrients at reference sites. The reviewer recommended a more traditional approach to summarize reference data, whereby each reference site's nitrogen and phosphorus observations are reduced to a site median, and then distribution statistics on the population of medians is calculated.

RESPONSE: Regarding the number of reference sites, MT DEQ believes it has a good reference site network and has been actively identifying and sampling reference sites for the past twelve years. (Limited work was also carried twenty years ago by Bahls et al. [1992].) From 2000 to 2009 much effort went towards identifying new sites. In some parts of the state (e.g., eastern Montana) staff has gone over the landscape several times and we are at the point where few if any new sites can readily be identified. As it stands, there are 185 different reference sites across the state and all major ecoregions are represented. Because of the relatively large overall number of sites, MT DEQ management indicated that the Reference Project should focus on resampling the network rather than seeking new sites. Some level III ecoregions (e.g., Idaho Batholith) would benefit from additional sites but it is unlikely that will occur in the near future.

We agree that the Brillouin Evenness Index formula should be provided in Update 1 with an explanation of why this approach was taken. This has been included in the final report. Regarding our use of the Brillouin Evenness Index vs. site medians to summarize the reference data, we offer the following. By taking the Brillouin Evenness Index approach, MT DEQ made the assumption that each nutrient observation in the dataset was independent even if collected from the same site. The vast majority of sample observations from the reference sites were collected a month apart, and MT DEQ has shown that such samples are usually temporally independent (Appendix A.3, Suplee and Sada de Suplee, 2011). We believe the data, after application of the evenness index to assure equitable representation of each site, provide a very valuable characterization of reference condition especially when a reader wants to know the true range of nutrient observations (minimum, maximum) in Montana reference sites during baseflow.

Stakeholders from the Montana Nutrient Work Group had earlier indicated that this was important to them. And as pointed out by one reviewer, the exact manner by which reference data are summarized is not terribly important because we do not carry out inferential statistics with the data, nor are criteria tied to a specific reference percentile. We agree with the reviewer that with the approach we used we cannot assure that there is no intra-site temporal pseudoreplication, an issue discussed at length in Hurlbert (1984). In response, we have now provided two summary statistics tables for each ecoregion; the original (derived, as before, using all observations and the Brillouin Evenness Index), and a 2nd table which shows the frequency distribution (25th through 90th) based on the median nutrient concentrations from each site. We believe this approach will provide readers the maximum amount of information and will make comparison to other work easier, since reduction of site data to medians is common in the literature (e.g., Robertson et al., 2001; Wang et al., 2007; Stevenson et al., 2012).

1.7. Concern that MT DEQ has Recommended Nutrient Criteria Concentrations in some Ecoregions Beyond the Applicable Reference Distribution

Two reviewers were concerned that nutrient criteria concentrations had been set at levels beyond any single observation collected in the regional reference streams. The Northern Rockies and Idaho Batholith are good examples. Although one reviewer agreed with MT DEQ that one should not use reference condition nutrient concentrations *alone* to set criteria, at the same time the idea of setting a criterion higher than the highest observation in the regional reference sites was clearly troubling to reviewers.

RESPONSE: The reviewers comments can be summarized as (1) concentrations beyond the reference distribution 75th percentile may be linked to known harm-to-use thresholds (e.g., via benthic algae density), but they will not be protective of sensitive, low-nutrient adapted organisms, and (2) in the ecoregions with naturally-low nutrients the harm-to-use concentrations derived from the dose response studies always had a range, and MT DEQ should have picked the lower concentration threshold given that we are operating beyond the bounds of the regional reference condition. Regarding point 1, we agree with the reviewer that if stream concentrations rise to the criteria and the criteria are beyond the reference condition, some organisms—like low-nutrient diatom taxa—would be displaced. The difficulty with establishing criteria to protect microscopic organisms like this is that there is no definitive harm to the beneficial uses established in Montana law. Studies generally show that with some additional nutrients ultra-oligotrophic streams will have more of the same macroivertebrates (as evidenced by O/E scores >1.0) and more robust populations of some fish. Fish—and to a somewhat lesser degree macroinvertebrates—link directly to Montana's beneficial uses. But as Montana state law is currently written, it would be difficult to defend a criterion based on protecting low-nutrient diatoms (as suggested by one reviewer).

Regarding the 2nd point, there is definitely merit to the idea that if there are several dose-response studies for an ecoregion and the concentrations from them generally fall beyond the reference distribution, greater weight should be given to the study or studies with the lower concentrations. As a result, in the final draft we have somewhat lowered the criteria recommendations in several ecoregions where this occurred. We have still kept an eye on maintaining the reference Redfield ratio, and in some cases the final concentrations are still beyond the reference distribution, but they are closer to it.

2.0 Selected Comments from Individual Reviewers

Here are important comments unique to individual reviewers.

2.2 How a Stream Reach is Delineated (Pertains to Assessment Methodology)

One reviewer was concerned that the flexible manner by which a stream reach can be delineated could make it difficult for any stream reach to ever be found impacted by nutrients, because data from impacted sites would be lumped with data from unimpacted sites and would, in effect, dilute the signal. The reviewer also noted that because of the flexibility in establishing assessment reaches, intentional manipulation of reach lengths could drive the outcome.

RESPONSE: The potential for unethical actions to manipulate analysis outcomes is always present in assessment work, but the high level of professionalism in the MT DEQ staff is such that this issue has not arisen. Regarding the flexibility of assessment reach lengths, this was done purposefully as discussed in detail in Appendix A.2.0 of Suplee and Sada de Suplee (2011). A basic assumption of the method is that reaches should be relatively homogenous in time (over the past 10 years) and in space, and observations collected within the reach should be largely independent. One reviewer was concerned about sample independence but MT DEQ has demonstrated independence in similar nutrient-concentration datasets using standard statistical tests (Durbin-Watson, Rank von Neumann). From these results and earlier experience, temporal and spatial independence guidelines were defined to make sure data collection maintains sample independence to the degree possible (nutrient and biological samples have to be collected a month apart at a site, for example).

If an assessor concludes that a reach is really *not* adequately homogenous (e.g., it comprises a distinctly impacted segment and an unimpacted segment[3]) it is incumbent upon the assessor to subdivide the reach and make an independent assessment of each new segment. This stratification allows maximal precision of estimates for minimal sampling effort (Norris et al., 1992). What remains constant is the minimum number of water quality and biological samples that need to be collected in each of these new assessment reaches in order to make a final compliance decision. The reviewer seemed to suggest that fixed reach lengths, numbers of sites, etc. along streams would be better, but experience has shown that this is highly impractical in applied assessment. If MT DEQ were to carryout assessments using fixed-length reaches, results would be far more arbitrary then the approach currently found in the SOP.

3.0 Summary of Changes Resulting from the Peer Review

1. We have included soluble nutrient data in the final report (Update 1) for the River Breaks.

2. The Brillouin Evenness Index formula is provided in Update 1 with a better explanation of why this approach was taken. We have also characterized reference using median datasets, further described below in 3.

3. Reference condition within an ecoregion has been characterized by first reducing data from each reference site to a site median, then calculating distribution statistics (25th, 50th, 75th, and 90th) for the ecoregion based on the population of site medians.

4. We agree that there is merit to the idea that if there are several dose-response studies for an ecoregion and the concentrations from them are generally beyond the reference distribution, greater weight should be given to the study or studies with the lower concentrations. As a result, in the final document we have somewhat lowered the criteria recommendations in several ecoregions where nutrient concentrations are naturally very low. We have still kept an eye on the Redfield ratio of the regional reference streams and, in some cases, the final criteria recommendations are still beyond the reference distribution, but they are closer to it.

4.0 References

Bahls, L., R. Bukantis, and S. Tralles, 1992. Benchmark Biology of Montana Reference Streams. Helena, MT: Montana Department of Health and Environmental Sciences.

Bothwell, M.L., 1989. Phosphorus-limited Growth Dynamics of Lotic Periphytic Diatom Communities: Areal Biomass and Cellular Growth Rate Responses. Canadian Journal of Fisheries and Aquatic Sciences 46: 1293-1301.

DEQ-7, 2012. Circular DEQ-7, Montana Numeric Water Quality Standards. October 2012. Helena MT: Montana Department of Environmental Quality.

---

[3] Temporal non-uniformity must also be considered. For example, if a new feedlot with several permit violations was built alongside the stream three years ago, the assessor would not be including data in the analysis collected six years ago.

Flynn, K., and M.W. Suplee, 2013. Using a Computer Water Quality Model to Derive Numeric Nutrient Criteria. Lower Yellowstone River, MT. WQPBDMSTECH-22. Helena, MT: Montana Department of Environmental Quality, 269 p plus appendices.

Horner, R.R., Welch, E.B., and R.B. Veenstra,1983. Periphyton of Freshwater Ecosytems. Wetzel,R.G. (ed.) Dr. W. Junk Publishers, The Hague.

Hurlbert, S.H., 1984. Pseudoreplication and the Design of Ecological Field Experiments. Ecological Monographs 54: 187-211.

Mapstone, B. D., 1995. Scalable Decision Rules for Environmental Impact Studies: Effect Size, Type I, and Type II Errors. Ecological Applications 5: 401-410.

Norris, R. H., E. P. McElravy, and V. H. Resh, 1992. "The Sampling Problem," in *The River Handbook*, Calow, P. and Petts, G. E., (Oxford, England: Blackwell Scientific Publications)

Rier, S.T., and R. J. Stevenson, 2006. Response of Periphytic Algae to Gradients in Nitrogen and Phosphorus in Streamside Mesocosms. Hydobiologia 561: 131-147.

Robertson, D.M., D. A. Saad, and A.M. Wieben, 2001. An Alternative Regionalization Scheme for Defining Nutrient Criteria for Rivers and Streams. U.S. Geological Survey, Water Resources Investigations Report 01-4073.

Stevenson, R.J., B.J. Bennett, D. N. Jordan, and R.D. French, 2012. Phosphorus Regulates Stream Injury by Filamentous Green Algae, DO, and pH with Thresholds in Responses. Hydrobiologia 695: 25-42.

Suplee, M.W., A. Varghese and J. Cleland, 2007. Developing Nutrient Criteria for Streams: An Evaluation of the Frequency Distribution Method. Journal of the American Water Resources Association 43: 453-472.

Suplee, M.W. V. Watson, A. Varghese, and J. Cleland, 2008. Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers. Helena, MT: Montana Department of Environmental Quality.

Suplee, M.W., and R. Sada de Suplee, 2011. Assessment Methodology for Determining Wadeable Stream Impairment Due to Excess Nitrogen and Phosphorus Levels. WQPBMASTR-01. Helena, MT: Montana Department of Environmental Quality. *Available at*: http://deq.mt.gov/wqinfo/qaprogram/sops.mcpx

Vitousek, P.M., J.D. Aber, R.W. Howarth, G.E. Likens, P.A. Matson, D.W. Schindler, W.H. Schlesinger, and D.G. Tilman, 1997. Human Alteration of the Global Nitrogen Cycle: Sources and Consequences. Ecological Applications 7: 737-750.

Wang, L., D.M. Robertson, and P.J. Garrison. 2007. Linkages Between Nutrients and Assemblages of Macroinvetebrates and Fish in Wadeable Streams: Implication to Nutrient Criteria Development. Environmental Management 39: 194-212.

Review of MT Nutrient Criteria Documents.

Response to Questions:

MT – Wadeable Streams Draft Peer Review Questions

1. MDEQ is considering two approaches for the derivation of numeric nutrient criteria in wadeable streams: (1) eco-regional reference condition, and (2) regional and non-regional stressor-response studies. Compare and contrast the ability of each approach to provide a sound scientific basis for numeric nutrient criteria derivation. Please provide documentation on any identified ranges protective of aquatic life based on similar studies. If possible, please provide alternate methodologies using available data and tools, and describe the corresponding advantages and disadvantages.

Many of my specific concerns are detailed separately, following my responses to #1-6, some of which address this question. My overall assessment of the MDEQ dual approach was that it was generally well thought out and appeared to have protection of streams in mind. There were a few exceptions where it appeared the criteria were set a bit too high (well beyond the 100% of reference distribution), and I commented on those decisions and provided citations where available. However, MDEQ did a commendable job of reviewing scientific literature and applying peer-reviewed literature in support of developing defensible, numeric criteria.

2. In Section 3.6.1., Montana suggests that no nutrient criteria are needed for streams in the 1 Level IV Ecoregion within the Northwestern Great Plains: River Breaks (43c). The MDEQ rationale for this decision is: "This level IV ecoregion has highly turbid, flashy streams with naturally elevated TP and TN levels. Concentrations observed in the region's reference sites indicate that nutrient concentrations here are already naturally elevated above the harm-to-use thresholds identified for the plains region as a whole. As such, no nutrient criteria are recommended for streams within this level IV ecoregion." Please comment on whether the state has provided a sufficient scientific basis that 1) these levels are naturally elevated, 2) additional increase in nutrients would not cause harm to aquatic life, and 3) that, therefore, criteria are not needed. Is the reviewer aware of any additional information that could be provided to either support the State's assessment of natural background or that could be used to derive site specific criteria?

I struggled with this decision. I cannot concur that additional increases in nutrients would not cause harm to aquatic life, mainly because there are not sufficient data to support this conclusion. I commented that there needs to be some consideration of dissolved nutrients, or at least a thorough discussion about them relative to TN and TP. Potential sources of nutrients to these streams are likely to be primarily dissolved in form and, without knowing whether there are high levels of dissolved N and P in reference sites, I cannot determine whether inputs are likely to harm aquatic life. Stream flow is also an important consideration. Extended periods of low flow during droughts coupled with over-enrichment from anthropogenic sources of N and P would likely result in biological responses that could harm aquatic life. I also noted that there was a very wide range of TN and TP values among different streams, suggesting that even level 4 ecoregions may not sufficiently capture the variability in geology and natural nutrient concentrations. In sum, I suggest this decision needs further consideration.

3. MDEQ is proposing to allow TN and TP criteria to be exceeded 20% of the time and be considered supporting aquatic life uses. This frequency was derived based on analysis of the Clark Fork River chl-a data. Please comment on the proposed exceedance frequency and whether allowing the stated magnitudes to be exceeded 20% of the time would not result in adverse effects on aquatic life. This information is discussed in the State's Assessment Methodology.

Allowing TN and TP criteria to be exceeded 20% of the time brings with it uncertainty about both the timing and the magnitude of exceedance. For example, exceeding 2 months in a row, in the middle of summer when flow is low is quite different than exceeding two distinct periods during a wet year with higher than average runoff. Similarly, exceeding a criterion by an order of magnitude just one time would obviously have different implications to aquatic life than if the criterion were exceeded by a few parts per billion.

The use of the Student's t-test to compare means to the criterion is MDEQ's approach to considering magnitude of exceedance. Although I am encouraged that MDEQ recognizes magnitude and frequency as important, I am not certain the t-test method is optimal. I outline my concerns with the sampling method, samples, and analysis of data for this test under point #6, below.

4. MDEQ's criteria approach includes a Chl-a value of 125 mg/m^2 to be used as part of the related assessment information. Please comment on the selection of chlorophyll as the primary response variable, the derivation of the chlorophyll threshold, and its application as a statewide assessment indicator.

Benthic CHLA is a widely used indicator of nutrient over-enrichment so it is defensible for MDEQ to include it as a measurement endpoint. However, benthic CHLA is not a reliable indicator of nutrient overenrichment because it is highly variable temporally due to periodic sloughing/senescing, grazing (Taylor et al. 2012) and scouring by high flows. In two years of sampling wadeable streams in central Texas, we found benthic CHLA to be one of the least reliable indicators of nutrient enrichment when compared to periphyton carbon: CHLA ratios, CNP ratios, enzyme activity, primary and bacterial production, and species composition (Scott et al. 2008, King et al. 2009, Scott et al. 2009, Lang et al. 2012). The observed frequency of exceeding 125 mg/cm2 CHLA could be highly variable depending upon the natural flow regime of a stream, interannual variability in precipitation, and timing of site visits, even though a stream may be vulnerable to dense periodic blooms that result in harm to aquatic life.

I also found the use of piecewise regression models (Dodds et al. 2002) to infer chlorophyll a values at particular nutrient levels to be questionable for a few reasons. It did not appear the confidence limits were considered. The fitted mean value falls within a highly variable cloud of points, indicating that 125 mg/m2 is exceeded in many streams possibly as much as half the time. If the goal is to keep CHLA below 125 mg/m2 a certain percentage of the time, quantile regression splines (Anderson et al. 2008) or other nonlinear quantile regression method would more closely match the objective. For example, if the goal was to keep CHLA < 125 mg/m2 80% of the time, the TP or TN value that aligns with the lower 5% CI of the 20% quantile would be a more appropriate number. Thus, risk of exceeding 125 mg/cm2 seems to be potentially high, or at least high uncertain, given the approach to derived the TN/TP thresholds and the high variability in benthic CHLA during the growing season.

5. Section 4.0 outlines a process for determining reach-specific nutrient criteria. Please comment on MDEQ's proposed approach for deriving reach-specific values.

The rationale and methods for setting criteria for this reach seem defensible. The process was consistent with the process used among ecoregions. Overall it is hard to find many suggestions on how they could improve their approach for setting criteria in these rivers, however see my previous comments about using CHLA as a biological endpoint, the use piecewise regression models to identify TP and TN criteria, and several point of concern about the sample design and statistical methods (see #6, below)..

6. Montana is proposing to interpret the numeric criteria using the Students t-test and binomial test to determine whether a stream segment is impaired. Please comment on the State's rationale for this approach.

It is obvious MDEQ has given this process a great deal of thought. Overall I am encouraged by the level of detail in the process and what appears to be a sincere attempt to develop criteria and a process for assessing criteria that is protective of aquatic life in the waters of Montana. This section is particularly important because it describes the nuts-and-bolts of how criteria are used to assess compliance.

There are several moving parts in this process that have the potential to strongly influence the outcome of an assessment. First, the manner in which reaches are delineated is flexible such that it seemed a bit ambiguous to me. Because sample "sites" allocated within reaches are used to assess criteria, how reaches are delineated could be manipulated to influence the outcome of assessments.

The use of multiple sampling sites within a reach to assess criteria is reasonable, but the scale of nutrient overenrichment required to fail a reach seems to be quite large. For example, under low, summer flow conditions, one site within a reach could conceivably fail during both visits whereas downstream reaches pass each time. The use of multiple downstream sites, some of which could be many kilometers away, to calculate exceedance frequency and mean nutrient levels ignores the local impairment and effectively "dilutes" the problem at this location, despite the fact it could span > 1 mile of stream (minimum distances between sites was 1 mile, correct?).

Another factor is the manner in which sampling locations and repeated measurements from those locations are used in the binomial test and t-test as if each sample unit reflects a measurement from the same population. There are two levels of organization being mixed here. Spatial and temporal sample units are being thrown in together in a haphazard way that ignores the distinct components of variance. If there were a clearer definition of reaches, site locations, and sample frequency from those sites, I would feel a little less uneasy about it, but as it stands, I get the impression that reaches may differ wildly in length, number of sites per reach will thus differ, and sample frequency may also differ.

The Clark Fork example illustrates the problem: 15-20 individual CHLA samples were collected per date and each "sample" was treated as a repeated measure, when in fact these are subsamples that are nested within a single observational unit (a site? I can't follow the sampling design very well). The total CHLA "samples" were 285-333 per site over a multiple-year period, but there were far fewer sampling events than 285-333, and far fewer TN and TP "samples" as well because those were composite grab samples. There also were different numbers of "samples" taken per site within the Clark Fork reach, as well as different numbers of samples within a site among dates. This type of analysis would not likely hold up in a peer-reviewed journal because each CHLA measurement is subsample of a single observational unit (site). In sum, I'm not necessarily saying that the approach will lead to wildly inaccurate assessments but I do believe that there are better ways to account for multiple measurements within a site and multiple dates per site within a reach to arrive at an estimate of exceedance frequency.

I also do not really see this as a hypothesis testing problem, but rather a risk assessment or probability of exceedance problem. There is a burgeoning literature on misuse of hypothesis testing statistics for ecological risk assessment and environmental assessment. The use of this approach for this particular application does not strike me as ideal.

I also am not certain about appropriateness of a t-test to detect magnitude of exceedance relative to the criterion. The t-test is a normal-distribution statistic that will be less likely to detect a difference in the mean relative to the criterion when data are skewed, and skewed data (infrequent but large departures from the criterion) are exactly why the statistic is being computed in the first place. Several other methods could be considered, ranging from computing empirical confidence limits using the bootstrap, to more sophisticated Bayesian approaches where an appropriate sample distribution is used and the test computes the probability that the sample mean differs from the criterion.

Specific comments, Addendum:

p2-2: Equitability of sample representation. I agree this is an important consideration but do not understand how the evenness statistic was applied to address the problem. How was J computed, specifically in terms of the observations in the nutrient database? The data are nested by sample unit (site),with each observation representing a distinct date, correct? More detail is needed here.

Section 2.5.1. This paragraph is interesting and I don't have any particular problems with the content except that it does not seem to have any direct applicability to criteria development in Montana. How was the information from sites that were intentionally enriched with N and/or P used to support criteria development? The section ends by suggesting this information was valuable for establishing a "lower bounds" for nutrient concentrations, but how was the information used? What are "lower bounds"?

Section 2.6. This section is an important addition to the document. I think the idea that differential nutrient limitation among different algal and other microbial species is not sufficiently acknowledged in the development in numerical nutrient criteria. This section does an excellent job of describing why managing for 2 nutrients is critical. However, I think a couple of ideas are used interchangeably and might need to be distinguished a bit.

The most important reason for differential nutrient limitation is that different species have different relative N and P demands thus one may be predominantly limiting to an aggregate endpoint such as benthic chlorophyll but in most circumstances at least some species are limited by another resource. This appears to be particularly true of photoautotrophs and heterotrophic microbes growing together in a periphyton community (Scott et al. 2008, 2009, Lang et al. 2012). In this paragraph, the idea of different species being limited by different nutrients is introduced, but later is conflated with the idea of communities switching back and forth between N and P limitation. These are 2 distinct ideas and should be parsed as such.

It is also unclear what is meant here by limitation. Limitation of accrual of benthic chlorophyll or something else? There are numerous indicators of limitation that may not manifest themselves as an increase in standing stocks if other factors are controlling accumulation in the short run. Enzyme activities, in particular, may reveal dual limitation of different subsets of species in the community whereas total biomass remains unchanged with enrichment of N, P or both because of the decoupling of heterotroph and autotroph recycling of carbon, N and P. I say this mainly to encourage a more explicit definition of limitation and acknowledgment that biomass accumulation may not be a good indicator of limitation in all situations.

The discussion about Redfield ratios is fine to include, but again it seems to be lumping responses into one large bin of either N or P limited, when in fact differential limitation means that each species in an attached community of photoautotrophs and heterotrophs has a different N and P demand, hence a community-level N:P ratio target is naive and potentially dangerous. I think ratios are a lot less important than concentrations and supply rate (velocity). Nutrient criteria should emphasize maintaining concentrations that fall below levels of individual nutrients that are known to overstimulate algae and/or microbes; the ratios at those levels may or may not be near "Redfield" because it is the supply rate of ions to the cells that ultimately determines whether a nutrient is limiting to growth or other physiological process.

In sum, I like the fact that Montana is thinking about these details but am a bit concerned about some of the overgeneralizations about nutrient limitation and nutrient ratios in driving decisions to manage for both N and P. The decision to manage for N and P need not be any more complicated than the fact that differential limitation probably occurs in most stream ecosystems and thus both nutrients are likely to limit some facet of the community at any point in time.

Section 3.0

I like the introduction to this section, detailing how the criteria are organized and presented in the forthcoming pages.

Fig 3-1 is a nice illustration of the distribution of reference sites. I noticed here and in the 2005 document that reference sites are spatially contagious. Large areas within each ecoregion are largely unrepresented by reference locations whereas other areas have high densities of them. This is a common problem, given that human activities tend to be clumped and thus the remaining "good" places are also clumped, away from human activity. However, given that there is some mention of the need for Level IV ecoregional criteria in some Level III ecoregions, it would be helpful to know whether there are some level IV ecoregions that contain few or no reference sites.

Fig 3-2. Red dots are cities? Not all red dots are labeled.

Section 3-1. Middle Rockies

Again, noting the Redfield ratio in the criteria recommendations. I don't think there is sufficient justification for including this number given it was derived for marine phytoplankton (i.e. is the the N:P ratio of marine phytoplankton). I worry about other states focusing on this ratio as they plod forward in their development of criteria. Also, the ratios reported are based on mass not moles so if ratios are to be reported please specify that they are based on mass.

p3-3, last paragraph: The interpretation of the breakpoint regression is correct, but more specifically the level of chla/m2 has reached its maximum (the bottom is effectively covered in filamentous algae). The first section of the breakpoint regression line is a quasi-linear increase with quite a bit of scatter. I don't like the interpretation of this type of regression because in reality what is happening is that the growth rate of Cladophora is faster at higher nutrient levels but with sufficient N and P will nevertheless grow until most of the channel is covered or until a high flow event knocks it back. The problem with assuming that a certain level of N or P will keep chla/m2 below a certain level is that it assumes that on average there are sufficiently frequent spates/high flow events that will keep the growth in check. In low water years or very dry summers I highly suspect that any level of N and P that is sufficient to promote filamentous algae will lead to unacceptable levels of chla/m2 (e.g. see experimental results in King et al. 2009). . If the goal is keeping chla/m2 below a certain level, other variables (particularly frequency/timing of storm events or high flows) are needed to better estimate the likelihood of failing to meet biological criteria. As currently written, I think it is overly simplisitic.

p. 3-4 Conclusions: The section acknowledges that TP as low as 20 is associated with undesirable outcomes. The use of N:P ratio as is further used to support 30 ug/L as a TP criterion because it maintains a 10:1 NP ratio, consistent with reference streams. Are we to presume that 200 ug/L TN is also associated with undesirable biological consequences as well? The justification for using the ratio as a basis for choosing 30 ug/L instead of 20 ug/L based on biological responses is warranted here. I feel there is too much emphasis on ratios without sufficient scientific documentation of it being as or more important than concentration/supply rate by ion. I am particularly concerned about the repeated reference to Redfield ratios.

Section 3.1.1 Level IV Ecoregion within the Middle Rockies: Absaroka-Gallatin Volcanic Mountains (17ia). There are only 4 reference sites in this region. The 4 sites span a huge range of TP, with as little as 16 ug/L. I find it hard to find support for a numerical criterion that would allow a stream with 16 ug/L TP to increase to 105 ug/L TP. I am confident there would be biological consequences. How realistic would it be to set basin-specific criteria for this subregion, given that it is relatively small?

Another concern is the selection of 250 ug/L TN despite the fact that this exceeds the highest reference site by almost 100 ug/L. It seems that given the high levels of TP that are naturally available in many of these streams, that any, small input of N could lead to nuisance growth of algae. In this region, it would be helpful to know the dissolved N levels because I suspect that most of the TN is particulate.. An addition of +100 ug/L NH4-N or NO3-N could lead to a substantial biological response.

3.2. Northern Rockies: Comments re: section 3.1 apply here as well.

3.3 Canadian Rockies: The very tight, extremely low TP values among all but one of the "reference" sites suggest that the selection of 25 ug/L TP for a criterion is too. high. It is far beyond the 75th percentile of reference as well as above the 20 ug/L TP number identified by other stressor response studies from the region. Again, I struggle with the use of explanatory models for predicting mean chla/m2. In most situations if TP is elevated it will be elevated by phosphate; adding 15+ ug/L TP above the highest reference sites has a high risk of impairing streams.

3.4. Idaho Batholith. Similar thoughts—TP is < 20 across all samples in reference sites. The literature review and discussion of previous results provides reasonable support for 30 ug/L TP, but not defensibly so. Setting the criterion at 30 ug/L seems to leave the door open for a minimum of 50% increase in P loads to these streams. Given this is far beyond the 75th and 90% reference site quantiles, I think greater justification is needed, especially considering the previous ecoregion was set at 25 ug/L TP despite similar reference distributions for TP.

Section 3.5. More detail on the sources of TN and TP in this region would be helpful. Is alder or another nitrogen fixing plant abundant in the uplands here? We see high natural concentrations of N in high alder streams in glaciated portions of Alaska but very low N when alder is low (Shaftel et al. 2012). As for P, is the source volcanic? What explains the high P levels in reference sites?

Also note that the discussion justifying the choice of criteria is long and somewhat speculative, although I appreciate the level of detail.

Section 3.6. The nutrient dosing study seems like it was not used directly supporting numerical criteria in this region beyond demonstrating that dissolved N and P additions stimulate algae. The amount of dissolved nutrients added was not particularly great despite the large algal and DO response, so it concerns me to see such high recommended levels of TN and TP for the region. However, the distribution of values among reference sites does support the recommended levels, assuming the reference sites are indeed representative of streams with minimal anthropogenic nutrient inputs. The large range of values among reference sites suggests that level IV ecoregions may be needed to parse out natural variability or there are streams that probably shouldn't be considered reference sites.

Section 3.6.1. River breaks. I follow the rationale for concluding that no criteria are necessary for this region. The lack of dissolved nutrient information makes it difficult to know whether all of the nutrients, particularly P, are bound to sediment or whether there is abundant dissolved N and P. I agree that dissolved nutrients can be variable due to biological uptake but in these systems I would suggest considering dissolved N and P. Without any criteria, it still seems these streams could be vulnerable to animal waste discharges, future wastewater discharges, or other sources likely to contribute very high levels of dissolved nutrients as well as organic matter. The streams are reportedly flashy, but this suggests there are periods of extended low flows between periodic flood events that permit blooms of phytoplankton and/or shallow wate r attached algae/plants. As explained in this document, I don't think the state has presented sufficient justification for electing not to set criteria for these streams.

Section 4. Reach specific criteria.

4.1 Flint Creek. The rationale and methods for setting criteria for this reach seem defensible.

4.2 Bozeman Creek et al. Overall it is hard to find many suggestions on how they could improve their approach for setting criteria in these rivers. The approach used is consistent with how it was done among ecoregions.

References

Anderson, MJ. 2008 Animal-sediment relationships re-visited: Characterising species' distributions along an environmental gradient using canonical analysis and quantile regression splines. Journal of Experimental Marine Biology and Ecology 366, 16–27

Dodds, Walter K., V. H. Smith, and K. Lohman. 2002. Nitrogen and Phosphorus Relationships to Benthic Algal Biomass in Temperate Streams. Canadian Journal of Fisheries and Aquatic Sciences. 59: 865-874.

King, RS, JM Taylor, JA Back, BA Fulton, BW Brooks. 2009. Linking observational and experimental approaches for the development of regional nutrient criteria for wadeable streams. Final Report. CP-966137-01. US Environmental Protection Agency Region 6, 151pp. http://www.baylor.edu/content/services/document.php/95606.pdf

Lang, D.A.*, R.S. King, and J.T. Scott. 2012. Divergent responses of biomass and enzyme activities suggest differential nutrient limitation in stream periphyton. Freshwater Science 31(4):in press.

Scott, J. T.*, D. A. Lang*, R. S. King, and R. D. Doyle. 2009. Nitrogen fixation and phosphatase activity in periphyton growing on nutrient diffusing substrata: Evidence for differential nutrient limitation in stream benthos. Journal of the North American Benthological Society 28:57-68

Scott. J. T.*, J. A. Back*, J. M. Taylor*, and R. S. King. 2008. Does nutrient enrichment decouple algal-bacterial production in periphyton? Journal of the North American Benthological Society 27:332-334.

Shaftel, R. S.*, R. S. King, and J. A. Back*. 2012. Alder cover drives nitrogen availability in Kenai Peninsula headwater streams, Alaska. Biogeochemistry 107:135-148

Taylor, JM, JA Back, RS King. 2012. Grazing minnows increase benthic autotrophy and enhance the response of periphyton elemental composition to experimental phosphorus additions. Freshwater Science 31 (2), 451-462

Review of Montana Department of Environmental Quality 2012 "Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers: Addendum 1"

General comments:

In preparation for the review below, I read Suplee et al. 2012 "Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers: Addendum 1", Suplee et al. 2011 "Assessment Methodology for Determining Wadeable Stream Impairment Due to Excess Nitrogen and Phosphorus Levels", and reviewed Suplee et al. 2005 " Identification and Assessment of Montana Reference Streams: A Follow-up and Expansion of the 1992 Benchmark Biology Study." In addition, I reviewed considerable literature to refresh my memory about details and look for additional information.

Overall, the "Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers: Addendum 1" was scholarly, thorough, and as scientifically sound as any state nutrient criteria document that I've reviewed. I think the approach is sound, information is relatively sufficient, and this work provides a very good next step for stakeholders of Montana and the development of their nutrient criteria to protect their resources. I have great respect for the originality and scientific rigor of the research conducted by MDEQ and its application in water policy. With that said, my responsibility is to indicate strengths and weaknesses in the approach and results, as well as address a set of specific questions. My review includes recommendations for additional approaches and sources of information for deriving benchmarks for nutrient criteria and selecting benchmarks for criteria that I hope MDEQ will find useful in revisions of this document or in their future work. Much research remains to refine the information needed for states and tribes to establish nutrient criteria that will adequately protect designated uses of their waters without overprotection. MDEQ is a leader in that effort and that effort serves the state of Montana well.

Below I've addressed the specific review questions and commented on related issues.

MT – Wadeable Streams Draft Peer Review Questions

1. **MDEQ is considering two approaches for the derivation of numeric nutrient criteria in wadeable streams: (1) eco-regional reference condition, and (2) regional and non-regional stressor-response studies. Compare and contrast the ability of each approach to provide a sound scientific basis for numeric nutrient criteria derivation. Please provide documentation on any identified ranges protective of aquatic life based on**

similar studies. If possible, please provide alternate methodologies using available data and tools, and describe the corresponding advantages and disadvantages.

MDEQ used ecoregion specific stressor-response relationships and ecoregional reference condition to derive numeric nutrient criteria for wadeable streams. Stressor-response relationships were used to determine the nutrient concentration at which undesirable effects in stream condition occurred. Ecoregional reference condition was used to determine the range of nutrient conditions at groups of sites with minimally impacted condition (*sensu* MDEQ 2005), that meet designated uses, and that have similar natural determinants of ecological condition (based on ecoregion constraints). Combining information from stressor-response relationships and ecoregional reference condition, nutrient criteria were then proposed for nutrient concentrations (both TP and TN) that were: 1) related to negative effects in biological condition that were predicted by stressor-response relationships and 2) greater than or equal to the 75th percentile of nutrient concentrations observed at reference sites. If sufficient knowledge is available for characterizing responses of valued ecological attributes (e.g. biological condition) to nutrient enrichment and minimally impacted nutrient concentrations at reference sites, and these characterizations are done appropriately, then I would argue that this is the best framework for deriving nutrient criteria. So an appropriate question to ask is, "Has MDEQ appropriately characterized nutrient concentrations in minimally impacted condition and responses of valued ecological attributes (e.g. biological condition and other indicators of designated use support) to nutrient enrichment?" I'll get to that question later after I briefly defend the MDEQ approach.

I have argued that nutrient criteria (and other stressor criteria) for a site should be derived with at least three steps (e.g. Stevenson et al. 2004, 2008; Soranno et al. 2008), given sufficient information:

1. determine expected conditions[1] for a site (which can be reference or desired conditions) based on management goals (which can be designated uses);
2. determine effect of nutrient concentrations on valued ecological attributes related to management goals for the site (e.g. biological condition or other indicators of designated uses) and select benchmarks in nutrient concentrations for possible criteria;
3. select benchmarks in nutrient concentrations that are greater than or equal to minimally disturbed condition and at concentrations with acceptable risk to impairment of valued ecological attributes (i.e. often measures or indicators of designated uses).

---

[1] Expected condition can be defined as minimally disturbed, least disturbed, best available, or desired condition (Stevenson et al. 2004). Here I use definitions of, least disturbed, best available from Stoddard et al. (2006) such that: minimally disturbed is "the condition of streams in the absence of significant human disturbance;" least disturbed is "found in conjunction with the best available physical, chemical, and biological habitat conditions given today's state of the landscape;" and best attainable condition is *"equivalent to the expected ecological condition of least-disturbed sites if the best possible management practices were in use for some period of time."* Desired condition is related to natural resources management and specifically addresses situations in which we management for attributes that may not be greatest in minimally disturbed conditions

In step 1 we should characterize the reference or desired (=expected) condition for the site that should include all physical, chemical and biological conditions that are related directly or indirectly to our management goals (e.g. designated uses) and that occur within the water, the riparian zone, the watershed, regionally, and even globally for contaminants transported through the atmosphere from distant sources. In some special cases, our goals may be to manage for desired condition (*sensu* Stevenson et al. 2004), such as more productive fisheries that are not characteristic of minimally disturbed conditions with high levels of biological condition (*sensu* Davies and Jackson 2006) in naturally low productivity ecosystems. Thus, tradeoffs between managing for productive fisheries and high levels of biological condition (biological integrity) are likely and should be addressed with tiered uses and different tiered uses for different waters within a region that meet the needs of regional stakeholders (Stevenson and Sabater 2010). Also, natural variation in climate, geology, hydrology, and water chemistry cause variation in minimally disturbed condition among ecoregions and among sites (e.g. Cao et al. 2007; Hawkins et al. 2010). So expected condition and nutrient critera, eventually, should be derived separately by ecoregion or by sites (e.g. Herlihy and Sifneos 2008; Soranno et al. 2008; Suplee et al. 2012 (the MDEQ document being reviewed)).

In step 2, we determine relationships between valued ecological attributes indicating designated and desired use support and nutrient concentrations. Nutrient concentrations are not a valued attribute because most people do not value them directly and only perceive risk from them if they cause problems to ecosystem services they do care about. There is little public support for managing nutrients independently of the effects that nutrients have on valued ecological attributes. We should not use reference condition nutrient concentrations alone to derive nutrient criteria because: 1) without stressor-response relationships we cannot be sure that nutrients affect valued attributes of the ecosystem and 2) we don't know the effects of incrementally increasing nutrient concentrations and at what nutrient concentrations risk of losing attributes become unacceptable. In evaluating stressor-response relationships, nutrient concentration benchmarks for potential criteria should be identified at the highest levels of nutrient concentrations at which an acceptable risk of losing valued attributes occurs. Thresholds in stressor-response relationships are highly valuable for delineating levels of nutrient concentrations at which risk levels change dramatically, thereby generating consensus among stakeholders for establishing criteria at specific nutrient concentration benchmarks.

In step 3, we determine what responses in valued ecological attributes change have acceptable risk benchmarks at nutrient concentrations greater than or equal to expected (usually reference) and then determine which benchmarks should be selected for nutrient criteria. In general, it's impractical (although not impossible) to manage a resource for nutrient concentrations lower than minimally or least disturbed condition, so nutrient criteria are usually at least as high as nutrient concentrations in reference conditions[2]; and criteria may be higher than reference conditions if valued attributes are not affected by nutrient concentrations less than or equal to reference conditions.

---

[2] Nutrient concentrations characteristic of reference conditions and supporting conditions of reference conditions are not any concentration within the range of nutrient concentrations observed at reference conditions. This will be discussed later in the text.

Now to the question, "Has MDEQ appropriately characterized nutrient concentrations in minimally impacted condition and responses of valued ecological attributes (e.g. biological condition and other indicators of designated use support) to nutrient enrichment?" Here I will also address elements of the review question:

- Please provide documentation on any identified ranges protective of aquatic life based on similar studies.
- If possible, please provide alternate methodologies using available data and tools, and describe the corresponding advantages and disadvantages.

I'll address the question and review question elements by criteria development step, and change the order of steps to correspond to the MDEQ methodology (characterizing stressor-response relationships and reference condition, and then deriving criteria).

Characterizing stressor-response relationships. MDEQ relies heavily on the relationships between nutrient concentrations and chlorophyll a, chlorophyll a and DO stress, and chlorophyll and aesthetics to related nutrient concentrations to support of designed uses. The nutrient-chlorophyll relationship is therefore the primary determinant of DO stress (e.g. Stevenson et al. 2012), which is an important stress on aquatic biota. The nutrient-chlorophyll relationship is also a primary determinant of aesthetics issues. Suplee et al. (2009) show reduced desirability of rivers for recreations use with chlorophyll a exceeding 125-150 mg chl a m^{-2}. The stressor-response relationships that they use are peer-reviewed and scientifically sound, or they have been developed by their own research in regions in which they have particular concern that that existing nutrient-response relationships would not apply. They consider different stressor-response relationships for different ecoregions, which is appropriate, because we would not expect high gradient streams, as in the mountains or foothills, to respond the same to nutrient pollution as in the low gradient streams of the prairies (see Stevenson et al. 2006 for example or ecoregion specific relationships). As an aside, I tried to compare the nutrient concentrations required to produce 125 mg chl a m^{-2}, but I could not determine which equation in Dodds et al. 2006 was equation 19. Comparing predicted nutrient concentrations at chlorophyll management targets using models in Mebane et al. (2009), Dodds et al. (1997 and 2006), and Stevenson et al. (2006) would be informative. Providing these models in the report would have been valuable for establishing the basis for the range in nutrient concentrations that were reported as required to maintain 125 mg chl a m^{-2}. Also, although results of experiments are based on soluble nutrients, Bothwell's experimental work with P and the N and P experimental work of Rier and Stevenson (2006) could be used to support determination of nutrient benchmarks for regulating chlorophyll a accrual.

While MDEQ's approach is scientifically sound, there are other relationships between nutrients and elements of stream ecosystems that may be important for determining whether nutrient pollution threatens designated uses of Montana waters. MDEQ definition of minimally impacted condition[3] indicates that more than chlorophyll and DO stress on invertebrates

---

[3] MDEQ (2005, p 2) defines minimally impacted condition as "Tier 2 — Minimally Impacted Condition" as "The characteristics of a waterbody in which the activities of man have made small changes that do not affect the completeness of the biotic community structure and function and the associated physical, chemical, and habitat conditions, and all numeric water quality standards are met and all beneficial uses are fully supported unless

should be included in stressor-response relationships. Since I did not find reference to the attributes specifically used to characterize designated uses of MT waters, I will mention some additional information that may be valuable to consider and which might not have been considered by MDEQ.

Relationships between nutrient concentrations, chlorophyll a, DO stress, and aesthetics likely cover most designated uses related to recreation, but may not protect biological condition of invertebrates, algae, and ecosystem function. Stevenson et al. (2008) observed very sensitive response of benthic diatom assemblages in the high gradient streams of the mid-Atlantic highlands with loss of sensitive species and deviations in species composition from reference condition at nutrient concentrations well below the 30 µg TP/L benchmark used for several MT ecoregions. With the abundance of periphyton data in the Western EMAP, the STAR reference site projects (Hawkins et al.), and now the National Rivers and Streams Assessment, generating informative stressor-response relationships for biological condition of periphyton and nutrients should be very practical.

In addition, Miltner and Rankin (1998), Yuan (2004), Smith et al. (2007), and Wang et al. (2007) describe invertebrate responses to nutrient concentrations that could be used to justify benchmarks for protecting biological condition of invertebrate communities. The mechanisms causing changes in species composition at relatively low nutrient concentrations are not well understood. DO and pH stress with nutrient enrichment are two likely mechanisms (Stevenson et al. 2012). In addition, release of streams and rivers from nutrient limitation enables invasion of habitats by taxa requiring higher productivity levels to survive and may shift competitive hierarchies in ways that cause loss of sensitive taxa adapted to naturally stressful low nutrient concentrations (Stevenson et al. 2008). Finally, release of aquatic ecosystems from nutrient limitation may enable invasion and reproduction of aquatic bacteria and fungi that could stress all other biota.

I applaude MDEQ's use of both TN and TP criteria because either can be limiting algal growth in streams with different geological conditions and resulting water chemistry, and at different times of years in some watersheds. I think this is largely done correctly given the amount of information available, where in high P reference regions MDEQ proposes low N criteria to constrain algal accrual. I think selected concentrations will be protective of high biomass in most cases where low N is used to constrain algal accrual. However, I do want to caution that we need to learn more to accurately quantify algal nutrient relationships with both TN and TP in the model, as was used by Dodds et al. (2002, 2006) and MDEQ. Such models violate Leibig's Law of the Minimum. MDEQ does address this in their report, but in reality, those justifications may not be sufficient. There is evidence in recent research that Leibig's Law of the Minimum does not hold, which makes me think algal biomass models with TN and TP linked are appropriate. Even though the science is a bit soft here, I would recommend using

---

measured impacts are clearly linked to a natural source. Minimally impacted conditions can be used to describe attainable biological, chemical, physical, and riparian habitat conditions for waterbodies with similar watershed characteristics within similar geographic regions and represent the water body's best potential condition."

the linked models and unlinked models as a multimodel approach for getting a range of conditions that would probably constrain algal biomass below the 125-165 μg chl a m^{-2} targets.

Characterizing reference condition. MDEQ's characterization of nutrient concentrations at reference sites suffered from a low sample size in three ways: 1) for all but a couple ecoregions, there were very small numbers of sites; 2) ror a couple ecoregions, there were fewer than 30 observations of nutrient concentrations at reference sites; and 3) repeated measures of nutrient concentrations at the same site are not independent. In the truest sense of pseudoreplication, the characterization of central tendency and variation in nutrient concentrations at reference sites suffers from some level of dependence in the samples.

The pseudoreplication issue should be addressed in a straightforward manner and put into a broader context so that it does not become overly important as a distraction from the relatively sound science that does underpin MDEQ's efforts. Although Suplee et al. (2011) address the pseudoreplication issue in another report, the key point is that it should be addressed. The broader context should include the following points. First, precise characterizations of percentiles are not that important because reference condition was used as a point of "reference" for nutrient benchmarks in stressor-response relationships where undesireable conditions developed. Second, the relative independence of repeated measures in reference condition is probably pretty low, given other sources of variability in estimates of nutrient concentrations in a stream: spatial and temporal variability in nutrient concentrations of streams and analytical error. I have argued this myself (Stevenson et al. 2006). However, repeated measures statistics can be calculated relatively easily to determine the relative dependence of measurements from the same site given overall variability and to correct estimates of variance among sites for dependency in repeated measures to more accurately characterize the central tendency and variation in nutrient concentrations at reference sites. The evenness approach (calculating evenness of measures among sites) that MDEQ uses is interesting, but it does not address pseudoreplication and dependent measurement issue directly.

Modeling expected nutrient concentrations at sites with land use-nutrient relationships is another method for characterizing the central tendency and variation in nutrient concentrations in minimally disturbed conditions. Modeling reference condition is valuable when the number of reference sites is low or quality of references sites varies between regions, which may have been the case in MT. Examples of different approaches for this kind of modeling can be found in Dodds and Oakes (2004), Herlihy and Sifneos (2008), Stevenson et al. (2008), and Soranno et al. (2008).

Typically, if an endpoint of management is used in criteria development, or as pseudocriteria, as chlorophyll a, then reference condition of that parameter is also described. Reference conditions were reported consistently for TP and TN concentrations. I'd recommend that chlorophyll a, diatom decreasers metric, and Hilsenhoff's biotic index (HBI) be described for reference conditions.

Selecting nutrient benchmarks for criteria. In general, if valued ecological attributes (direct indicators of designated use support) respond sensitively within the range of nutrient conditions at reference conditions, it is difficult to justify higher nutrient benchmarks than the

75th percentile of reference condition, assuming reference condition supports designated uses as described by MDEQ. I remember three distinct exceptions to this rule in MDEQ's proposed criteria. One is several Rockies ecoregions, in which proposed nutrient criteria were substantially above background concentrations, the other was in an ecoregion in which P was high and the TN criterion was well above the 75th percentile of reference condition, and the other was in the River Breaks region in which no criteria were proposed because no known ecological responses to nutrients were known for concentrations that high. I'll address the River Breaks situation below with the specific question asked for the review.

I'm concerned about selecting nutrient criteria above background concentrations in the Rockies ecoregions because proposed criteria would not protect sensitive, low nutrient diatom taxa, ecosystem functions of low productivity systems, and likely corresponding biodiversity of other groups whose response to low nutrient concentrations are poorly understood (bacteria, meiofauna, even benthic macroinvertebrates species). In Stevenson et al. (2008) we observed substantial changes in species composition of diatom assemblages at low nutrient concentrations and substantial loss of sensitive, low nutrient taxa (from counts) across the range of nutrient conditions. I've seen the loss of sensitive, low nutrient taxa (from counts) with low levels of nutrient enrichment in the extensive ecological assessment work that I've done around the country. Yes, this is just loss of taxa from counts, and we're not quite sure what that means (although my students and I are trying to understand that more), but we may actually be losing more taxa from the habitat (not just counts), as well as losing fewer. At this point, it just depends upon the weight of assumptions in the model. But if this is true for diatoms, then what about other groups. Also allowing higher N concentrations as well as P concentrations could impair biological integrity of these minimally disturbed, near-natural systems. For example, releasing N limited systems from severe N limitation could cause loss of diatoms with N-fixing cyanobacterial endosymbionts (e.g. *Epithemia*) or allow invasion of potentially nuisance taxa. Also, the relaxed P and N criteria are close to thresholds for releasing systems from severe nutrient constraint, so nuisance growths of algae could occur more frequently than if nutrient criteria were constrained to reference condition. Quantifying acceptable risk of nuisance growths should guide considerations.

Setting stressor criteria at a stressor level predicted to cause a target responses (i.e. nutrient criteria at nutrient concentrations at which a model predicts a target 125 mg chl a m^{-2}) means that when the stressor is at that level, the response with be greater than the target 50% of the time and less than the target by 50% of the time and by a magnitude that is related to the mean square error of the predicted values. Should quantile regression or conditional probabilities (Paul and McDonald 2005) be used to determine the stressor level that will manage the response with an acceptable frequency and intensity of exceedance?

Thresholds, relatively abrupt changes in rates of response along stressor gradients, are valuable for deriving environmental criteria (Muradian 2001). They identify benchmarks for possible nutrient criteria and help determine which benchmarks should be used as criteria. Some threshold responses are more valuable than others (Stevenson et al. 2008). Information for different threshold responses should be interpreted differently. For example, a response showing assimilative capacity and then a threshold response as stressors increase (Stevenson et al. 2008, Figures 2A & B, , sometimes called a type III response (http://en.wikipedia.org/wiki/Functional_response)) is particularly valuable for deriving criteria because stressor levels just below the threshold are clearly protective of reference condition and provide a margin of safety. Thresholds in responses showing high rates of change at low stressor levels and little change at high stressor levels, sometimes called a saturation curve or type II response (Stevenson et al. 2008, Figure 2C, (http://en.wikipedia.org/wiki/Functional_response), is more difficult to apply in criteria development. Nutrient uptake and algal growth have this type response along nutrient gradients with highly sensitive responses to nutrients at low concentrations and little response to nutrients at high concentrations.

The question then becomes, "How low should you set criteria to constrain growth in a part of the curve that is relatively linear?" Setting criteria just below the threshold (or breakpoint as described in Dodds et al. 2002, 2006 and as applied in this MDEQ document) provides little protection from adverse effects. Algal growth and accrual are largely at their greatest levels at nutrient concentrations just below those breakpoints. So in Transitional Level IV Ecoregions of the Northwestern Glaciated Plains, a justification for TN criteria of 560 µg/L was 560 was lower than the maximum saturation threshold (with saturation thresholds of 367 and 602), this criterion would not constrain algal accrual if these models are correct (which is the assumption of using them). The real explanation for choosing that level would seem to be that 560 µg TN/L is close to the 75th percentile of reference condition and you can't expect to do much better than that, even though biomass-nutrient models indicate biomass accrual could be near maximum levels at that TN concentration.

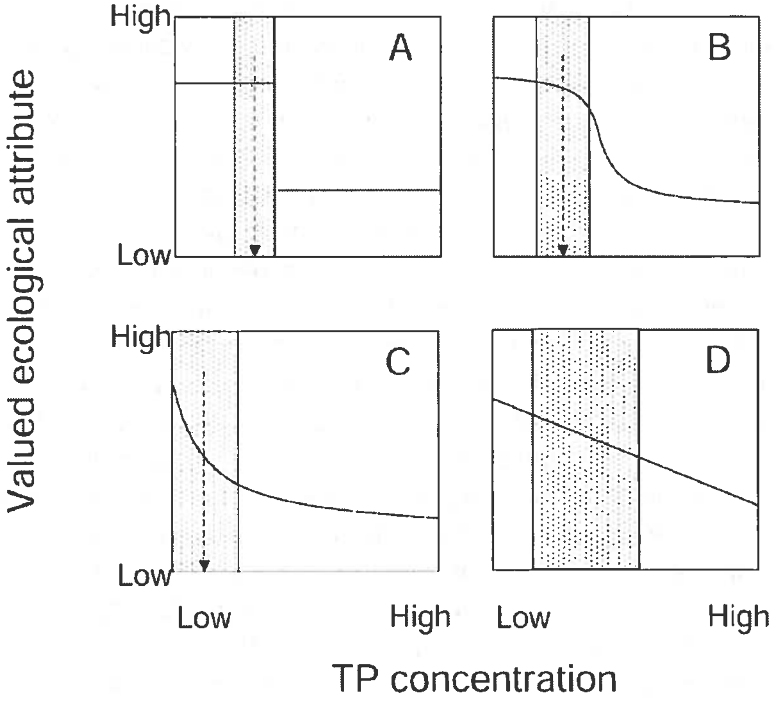

FIG. 2. Approaches to development of stressor criteria when potential responses of valued ecological attributes to stressors (e.g., total P [TP]) are nonlinear with assimilative capacity for increases at low levels the stressor (A, B), nonlinear with strong sensitivity to changes at low levels of the stressor (C), and linear (D). A stressor criterion is established at a level of a stressor that protects the valued ecological attribute. Arrows indicate TP criteria justified on the basis of the form of the stressor–response relationship. Shaded areas indicate the range of TP criteria that could be acceptable. Acceptable ranges vary as a function of the linearity of the stressor–response relationship and the type of nonlinear relationship.

Implementation. I like the concept of Tier I and II assessments for determining whether sites meet nutrient criteria. This does relate to an issue about risk of use support and the way we use statistics to define reference condition and determine whether a site meets its water quality criteria. I'll discuss this under a later question about exceedance frequencies. In particular, I like the use of biological condition assessment in the Tier II assessments with the diatom and macroinvertebrate metrics. These metrics should provide a temporally integrated signal that should complement the temporally variable assessments of nutrient concentrations. I would, however, recommend that MDEQ include metrics that evaluate decrease in sensitive native taxa as well as the increase nutrient pollution tolerant taxa (e.g. diatom increases and HBI taxa), because these taxa are key elements of biological condition for which we manage waters (Davies and Jackson 2006).

MDEQ chose to implement criteria during the growing season only, which assumes that the mechanism by which nutrients affect designated uses is by stimulating algal growth and that recreational use exposure is during the growing season. This is likely true, that algae do not bloom to nuisance levels or threaten low DO or high pH during non-growing seasons. But there are other potential ways that nutrient pollution can affect aquatic life use, which are poorly understood and poorly documented (i.e. shifts in competitive hierarchies and disease), and some nuisance growths of diatoms that alter habitat structure can occur during cooler seasons of the year.

I'm surprised there is little difference in when the growing season occurs. Why not use water temperature (for algal endpoints) and degree days (for invertebrate endpoints)? Do these time periods allow for interannual variation?

Tiered aquatic life uses should be considered (Davies and Jackson 2006). The problem with potential management challenges in the Rockies (as well as elsewhere), where reference nutrient conditions seem really low and well below most targets for designated use, plus the desire for P enrichment to support fisheries and limit *Didymosphaenia* blooms, is that we could lose attributes of natural systems that now exist. Tiered aquatic life use policies could allow protection of some systems within those ecoregions for near natural structure and function (which now exist) and allow other systems to be managed for fisheries and *Didymo* control.

I sense several issues touch on the policy doctrine of "independent applicability" of stressor and response criteria. Conceptually, one reason to set criteria within reference conditions (with a margin of safety) is because there may be negative responses that we don't know about if stressors are higher. In a perfect world, we know all the possible responses to stressors, so we could relax stressor criteria to levels that protect desired responses with acceptable risk. BUT, do we know enough about nutrient effects on designated uses to make relax criteria to ranges outside the reference condition (i.e. greater than the 75th percentile of reference condition as argued below)? MDEQ does use elements of independent applicability in their assessments. For example, level I assessments only involve comparisons of nutrients and not biological endpoints to nutrient criteria. In addition, if either N or P fail, then the system is not in compliance (MDEQ 2011, pp. 3 and 4), both do not need to fail to be in noncompliance. I did not find and review the assessment methodologies for level II decisions with sufficient detail to evaluate issues related to independent applicability.

MDEQ is as knowledgeable about potential nutrient impacts as any other state or tribal agency. They have chosen a level of risk with which they are comfortable for protecting their waters. Rivers are relatively resistant ecosystems. If errors are made, given the MDEQ "good-faith" effort, designated uses of the rivers should be able to be restored, unless there is regional extirpation of taxa which is unlikely in the short term.

2. **In Section 3.6.1., Montana suggests that no nutrient criteria are needed for streams in the 1 Level IV Ecoregion within the Northwestern Great Plains: River Breaks (43c). The MDEQ rationale for this decision is: "This level IV ecoregion has highly turbid, flashy streams with naturally elevated TP and TN levels. Concentrations observed in the region's reference sites indicate that nutrient concentrations here are already naturally elevated above the harm-to-use thresholds identified for the plains region as a whole. As such, no nutrient criteria are recommended for streams within this level IV ecoregion." Please comment on whether the state has provided a sufficient scientific basis that 1) these levels are naturally elevated, 2) additional increase in nutrients would not cause harm to aquatic life, and 3) that, therefore, criteria are not needed. Is the reviewer aware of any additional information that could be provided to either support the State's assessment of natural background or that could be used to derive site specific criteria?**

I don't like the idea that there are no nutrient criteria set for waters, even given the rationale that natural concentrations are naturally high and no instream or downstream effects are expected to occur. It makes me nervous that we know enough about nutrient-stream relationships to make that call. Will antidegradation policy prevent this system from getting worse? Why not have criteria be existing condition, i.e. the reference condition, as the criterion? Independent applicability would call for using reference condition of a contaminant in this case. Addressing questions 1-3. 1) I am not convinced that MDEQ has provided a sufficient scientific basis that nutrient concentrations are naturally elevated because: the quality of reference sites relative to land use in this region is not described, so how minimally disturbed is the reference condition; how is minimally disturbed and meeting designated used defined in this ecoregion if the systems are so naturally stressed; the number of reference streams sampled is low (n=8), even though the number of samples in relatively high (n=29), but note the 3 outlying samples with TP > 3.6^{10} (i.e. >3981 µg TP/L) that are likely from the same stream and indicating a site-specific dependence; and modeling reference condition with nutrient and land use data from all sites in the region may help better evaluate minimally disturbed conditions. 2) It does not seem likely that there are no instream or downstream effects of elevated nutrients because: phytoplankton blooms can occur during storm-free periods when waters slow and clear ; and downstream effects seem likely because patches of this ecoregion are so small and waters having to flow somewhere. 3) Criteria should be established to prevent dumping in this region, prevent degradation, and prevent surprises.

4. **MDEQ's criteria approach includes a Chl-a value of 125 mg/m^2 to be used as part of the related assessment information. Please comment on the selection of chlorophyll as the**

primary response variable, the derivation of the chlorophyll threshold, and its application as a statewide assessment indicator.

Using chlorophyll a or any indicator of valued ecological attributes, such as the diatom decreasers and Hilsenhoff biotic index, is an important check an assessments based on stressors because they directly address whether uses are being met. Chlorophyll a is a particularly important variable to use in determination of nutrient criteria and assessments of site compliance because it is probably the best indicator of algal biomass that we have and most effects of nutrients on designated uses of rivers and streams are caused by stimulation of either benthic or planktonic algal growth. MDEQ's derivation of 125 mg chl a m^{-2} as a management target to protect recreational use of rivers and aquatic life from DO stress is a model for what should be done by other states and tribes. In general, the chlorophyll standard was appropriately varied from region to region when reference condition nutrients were in the 30 µg TP/L and 300 µg TN/L range, but I do have concerns about using chlorophyll as an endpoint in the Rockies ecoregions where reference nutrient concentrations are low and nuisance levels of chlorophyll causing impairment of aesthetics and DO are not the only likely cause of changes in biological condition. Protecting biological condition at near natural levels may, however, be above the level of protection that stakeholders support in Montana. Although, tiered uses or an outstanding resource waters protection could be used to protect at least some low nutrient systems from increased productivity and resulting changes in biological condition. Other than these overall comments, details supporting the comment for question 4 are covered under question 1.

5. **Section 4.0 outlines a process for determining reach-specific nutrient criteria. Please comment on MDEQ's proposed approach for deriving reach-specific values.**

There are special situations when establishing nutrient criteria based on regional reference condition may be too high or too low. In the case of the Georgetown Lake Dam, the state statues call for a recalibration because they won't alter the location of the intact. The flow weighted approach in Bozeman Creek, Hyalite Creek, and East Gallatin also seems sound. I do question the relaxation of TN criteria above the very low 100 µg/L reference condition to around 250 µg TN/L, again for protecting high quality waters in Bozeman Creek, Hyalite Creek, and East Gallatin. This is the same issue as discussed for very low TP conditions in many of the Rockies ecoregions, but protecting high levels of biological conditions is a different issue than whether reach-specific criteria were determined appropriately based on management endpoints related to algal biomass, DO stress, and aesthetics.

3. **MDEQ is proposing to allow TN and TP criteria to be exceeded 20% of the time and be considered supporting aquatic life uses. This frequency was derived based on analysis of the Clark Fork River chl-a data. Please comment on the proposed exceedance frequency and whether allowing the stated magnitudes to be exceeded 20% of the time would not result in adverse effects on aquatic life. This information is discussed in the State's Assessment Methodology.**

6. **Montana is proposing to interpret the numeric criteria using the Students t-test and binomial test to determine whether a stream segment is impaired. Please comment on the State's rationale for this approach.**

I want to address questions 3 and 6 together. I think they are related. They are kind-of statistical issues.

First, I'd expect that 20% or more of observed TP and TN conditions would exceed criteria levels at sites maintaining an average target condition of, for example 150 mg chl a m^{-2}. The way that the nutrient criteria have been developed is based on the relationships between nutrient concentrations and chlorophyll observed at a site (Dodds et al. 1997, 2002, 2006). I'm going to use one of my Figures from Stevenson et al. (2006, redrawn and rescaled) to illustrate this because the non-linear relationships illustrated in the Dodds et al. (1997) paper are too complex to illustrate these principles and the plotted data looks a bit off in Dodds et al. (2002), which may be related to the erratum of Dodds et al. (2006).

In the statistical models that we use, there is both variation in the measurements of the independent and dependent variables (see Figure 1, blue and red lines respectively could be standard error bars of predicted and measured values). For example, when the average nutrient concentration is 30 µg TP/L in Michigan streams, chl a is approximately 20 mg/m^2. Michigan streams are grazer dominated and little periphyton accumulates with increases in nutrients, unless *Cladophora* can escape grazer control. Note values in Michigan are much lower than Kentucky by almost an order of magnitude. MDEQ based nutrient criteria on algal-nutrient model predictions to maintain biomass at a specific level or lower – usually 125 mg chl a m^{-2} target. So we should expect that nutrient concentrations will sometimes exceed the 125 mg chl a m^{-2} concentrations of the model, because there is variation around the predicted value. Actually, I'd expect the exceedance frequency to approach 50% as observed average nutrient concentrations at a site approach the criterion.

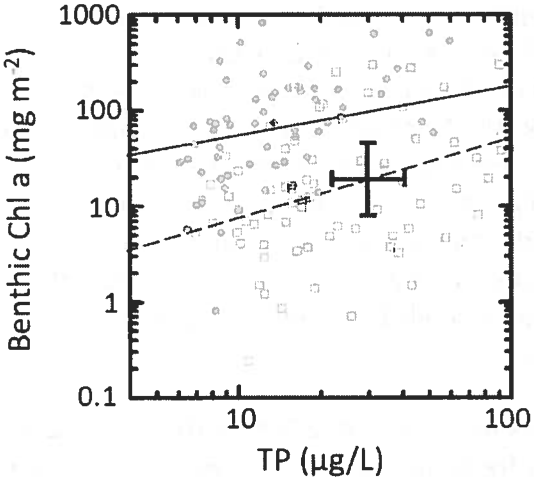

Figure 1. Figure 5 from Stevenson et al. (2006) redrawn and rescale. Open squares are values for streams in Michigan and shaded circles represent streams in Kentucky. Lines, dashed and solid respectively, are the linear relationships among the points.

This increase in exceedance frequency with nutrient concentration and algal biomass is illustrated in Suplee et al. (2011) Figure A4-1. Even for sites with the three lowest nutrient exceedance frequencies (and three lowest nutrient concentrations), maximum summer chl a is frequently greater than the 150 mg/m^2 expectation for the Clark Fork. The exceedance frequency actually provides a measure of risk of losing an attribute, in this case it's an aesthetically pleasing recreational venue and the potential for a DO event, which depending upon severity and extent, could have long-term repercussions for some biota. MDEQ have provided a margin of safety with lower biomass targets than impair aesthetics or cause DO stress and often lower nutrient creiteria than predicted to generate the target chlorophyll concentrations. This margin of safety may be the reason for average exceedance frequency at sites meeting uses being less than 50% even with nutrients concentrations near criteria.

Finally, I address the t-test and binomial test issues and issues with frequency distributions based on observations versus means. These issues are related to the risk of use support being affected by the way we use statistics to define reference condition and determine whether a site complies with water quality criteria. I want to start this discussion by reviewing a rationale for using frequency distributions of observations from reference conditions and a mean from a test sites to assess compliance at the test site. Then I'll transfer those concepts to evaluate how MDEQ's approach affects risk of supporting designated uses based on using observations, means, regression, binomial tests, and t-tests.

Consider the following scenario. Reference sites are selected because they are minimally disturbed based on land use and they support a specific level of aquatic life (and/or other uses) with an acceptable risk (let's say a management endpoint like chlorophyll a exceeds criteria 10% of the time). A frequency distribution is used to characterize central tendency and variation in a stressor (e.g. nutrients) that affect designated use within the range of conditions at reference sites. The frequency distribution is based on single samples from reference sites within an ecoregion (single independent observations in a statistical sense, Y_i, Figure 2A). We then use the 75th percentile of the frequency distribution of observed stressor conditions at references sites (Y_i^{75}) as a criterion, because we recognize that conditions vary around the average or median condition for reference sites due to spatial and temporal variation related to weather, flow, time of day, etc. and measurement error. Conceptually we use the 75th percentile of the frequency distribution of observed stressor conditions at references sites (Y_i^{75}) as a criterion because we feel that it's protective. Why? Well, one reason may be the statistical rule related to hypothesis testing called the 75% error bound. This is a rule of thumb that you can use to compare two means such that: if the mean of one sample (i.e. group) of observations (test sites) is outside the 75th confidence interval of the other sample of observations (reference sites), you can assume that there is little

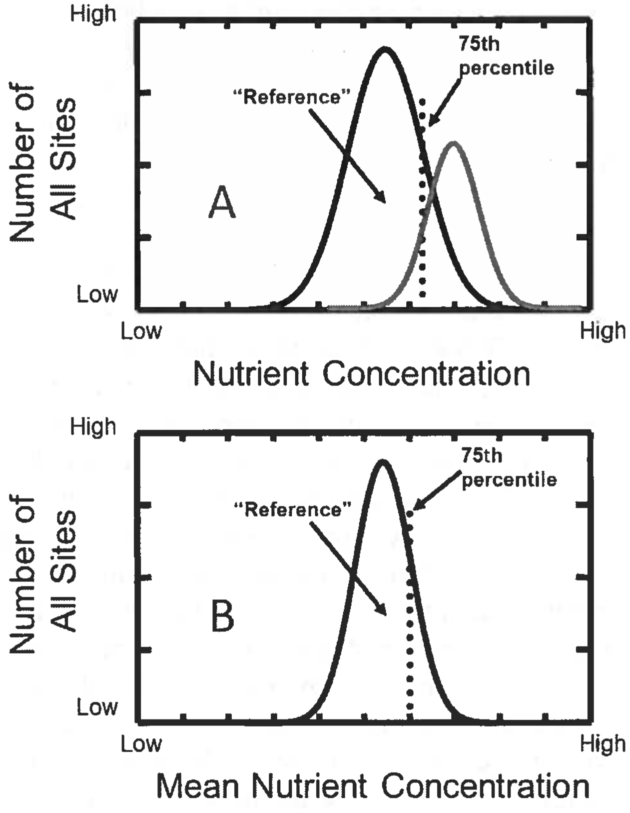

Figure 2. Distribution of nutrient concentrations when concentrations are represented by single samples site (A) and when concentrations are represented by means of samples at a site (B). The black distributions are reference site distributions. The blue distribution is the distribution of observations from a test site that is theoretically greater than the 75th percentile of the reference distribution according to a t-test or a binomial test.

probability that means of the first and second samples (groups of test and reference sites) are equal. Basically, 0.25 (1-0.75) is the attained significance for the difference between two means if the sample size is just 2 (if I remember correctly). So the idea is that conditions in the test set of sites would be different than the reference sites if the mean of test sites was greater than the 75th percentile of reference sites (based on single observations per site, or multiple observations from a smaller set of site that we could assume were independent). If agencies based development of criteria on mean measurements from a site, then the variance of the mean (Figure 2B) is much smaller than the variance of observations (Figure 2A). Comparing mean conditions at a site to the 75th percentile of a frequency distribution of mean observations at a site would be overprotective. Testing that mean of the test sites is significantly greater than the 75th percentile of reference sites (blue distribution in Figure 2A) versus just significantly greater than the mean of the reference site, would be underprotective.

MDEQ has proposed nutrient criteria that are the 75th percentile of reference condition or a higher concentration that is predicted to produce an effect that is undesireable. If my understanding of this process is correct, then using a t-test to determine whether mean conditions are greater than nutrient criteria would be underprotective, i.e. exceedance frequencies would be very high at sites before a site was found to be noncompliant. The mean concentration at the test site would have to be greater than the 75th percentile of the reference condition or the predicted level of nutrients causing a problem by an amount related to the variance in observed nutrient concentrations at the test site, the number of samples from the test site (n), and a t-statistic (which has a value of 2 when n is high). Issues associated with a binomial test are similar – some proportion of observed test site nutrient concentrations greater than 50% have to be greater than the criterion that is set at the maximum concentration that protects designated uses.

To counter these statistical issues with use of a t-stat causing underprotection of test sites, MDEQ does seem to have employed some margin of safety in setting the criteria at 125 mg chl a m^{-2}, which is below the 150 mg maximum okay level. In addition, MDEQ has adjusted acceptable levels of type I and II errors to reduce the problem of not detecting problems when they exist, and MDEQ is using chl a, a diatom indicator, and an HBI response criteria when nutrient concentrations are not obviously high. These additional rules can generate either greater under- or over-protection of waters. If only one the five criteria (N, P, chl, diatoms, HBI) has to fail, then that makes the assessment of compliance more protective than if the two nutrient criteria or all nutrient and biocriteria have to fail. I was not able to find details about compliance rules, which are apparently embedded in the spreadsheet that is referred to, so I can't evaluat Tier II assessments later. Tier II assessments, in which additional samples are collected and information is gathered also improves detection of non-compliance, reducing the error variation around means for the test site, which reduces possible difference between means at the test site and the nutrient criterion.

Overall, if I were a stakeholder concerned about protecting valued attributes within the streams, I'd be more concerned about potentially high risk of frequent loss of valued conditions (exceedance frequencies) when criteria are set at stressor-response model predictions that are too close to unacceptable levels of conditions, not the 20% exceedance problem or the t-test issue. The next frontier in deriving nutrient criteria may be bringing in a stronger risk assessment (e.g. Paul and McDonald 2005).

References

Cao, Y., C. P. Hawkins, J. Olson, and M. A. Kosterman. 2007. Modeling natural environmental gradients improves the accuracy and precision of diatom-based indicators. Journal of the North American Benthological **26**:566-585.

Davies, S. P., and S. K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16:1251-1266.

Dodds, W. K., V. H. Smith, and B. Zander. 1997. Developing nutrient targets to control benthic chlorophyll levels in streams: A case study of the Clark Fork River. Water Research 31:1738-1750.

Dodds, W. K., V. H. Smith, and K. Lohman. 2002. Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. Canadian Journal of Fisheries and Aquatic Sciences 59:865-874.

Dodds, W. K., and R. M. Oakes. 2004. A technique for establishing reference nutrient concentrations across watersheds affected by humans. Limnology and Oceanography: Methods 2:331-341.

Dodds, W. K., V. H. Smith, and K. Lohman. 2006. Erratum: Nitrogen and phosphorus relationships to benthic algal biomass in temperate streams. Canadian Journal of Fisheries and Aquatic Science 59:865-874.

Dodds, W. K., W. H. Clements, K. Gido, R. H. Hilderbrand, and R. S. King. 2010. Thresholds, breakpoints, and nonlinearity in freshwaters as related to management. Journal of the North American Benthological Society 29:988-997.

Hawkins, C. J., Y. Cao, and B. Rober. 2010. Method of predicting reference condition biota affects the performance and interpretation of ecological indices. Freshwater Biology **55**:1066-1085.

Herlihy A.T. and J.C. Sifneos. 2008. Developing nutrient criteria and classification schemes for wadeable streams in the conterminous US. Journal of the North American Benthological Society 27:932-948.

Muradian, R. 2001. Ecological thresholds: a survey. Ecological Economics 38:7-24.

Paul, J. F., and M. E. McDonald. 2006. Development of empirical, geographically specific water quality criteria: a confidential probability analysis approach. Journal of the American Water Resources Association 41:1211-1223.

Smith, A. J., R. W. Bode, and G. S. Kleppel. 2007. A nutrient biotic index (NBI) for use with benthic macroinvertebrate communities. Ecological Indicators **7**:371-386.

Soranno, P.A., K.S. Cheruvelil, R.J. Stevenson, S.L. Rollins, S.W. Holden, S. Heaton, and E. Torng. 2008. A framework for developing ecosystem-specific nutrient criteria: Integrating biological thresholds with predictive modeling. Limnology and Oceanography 53:773-787.

Stevenson, R.J. and S. Sabater. 2010. Understanding effects of global change on river ecosystems: science to support policy in a changing world. Hydrobiologia 657:3-18.

Stevenson, R. J., B. C. Bailey, M. C. Harass, C. P. Hawkins, J. Alba-Tercedor, C. Couch, S. Dyer, F. A. Fulk, J. M. Harrington, C. T. Hunsaker, and R. K. Johnson. 2004. Designing data collection for ecological assessments. In: M. T. Barbour, S. B. Norton, H. R. Preston, and K. W. Thornton, eds. Ecological Assessment of Aquatic Resources: Linking Science to

Decision-Making. Pgs 55-84. Society of Environmental Toxicology and Chemistry, Pensacola, Florida. ISBN 1-880611-56-2.

Stevenson, R. J., B. C. Bailey, M. C. Harass, C. P. Hawkins, J. Alba-Tercedor, C. Couch, S. Dyer, F. A. Fulk, J. M. Harrington, C. T. Hunsaker, and R. K. Johnson. 2004. Interpreting results of ecological assessments. In: M. T. Barbour, S. B. Norton, H. R. Preston, and K. W. Thornton, eds. Ecological Assessment of Aquatic Resources: Linking Science to Decision-Making. Pgs 85-111. Society of Environmental Toxicology and Chemistry, Pensacola, Florida. ISBN 1-880611-56-2.

Stevenson, R.J., S.T. Rier, C.M. Riseng, R.E. Schultz, and M.J. Wiley. 2006. Comparing effects of nutrients on algal biomass in streams in 2 regions with different disturbance regimes and with applications for developing nutrient criteria. Hydrobiologia 561:149-165.

Stevenson, R.J., B.E. Hill, A.T. Herlihy, L.L. Yuan, and S.B. Norton. 2008. Algal-P relationships, thresholds, and frequency distributions guide nutrient criterion development. Journal of the North American Benthological Society 27:783-799.

Stevenson, R. J., B. J. Bennett, D. N. Jordan, and R. D. French. 2012. Phosphorus regulates stream injury by filamentous algae, DO, and pH with thresholds in responses. Hydrobiologia 695:25-42.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16:1267-1276.

Suplee, M.W. and V. Watson. 2012. Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana's Wadeable Streams and Rivers—Addendum 1. Helena, MT: Montana Dept. of Environmental Quality.

Suplee, M.W., and R. Sada de Suplee. 2011 Assessment Methodology for Determining Wadeable Stream Impairment Due to Excess Nitrogen and Phosphorus Levels. Helena, MT: Montana Dept. of Environmental Quality.

Suplee, M.W., R. Sada de Suplee, D. Feldman and T. Laidlaw. 2005. Identification and Assessment of Montana Reference Streams: A Follow-up and Expansion of the 1992 Benchmark Biology Study. Helena, MT: Montana Dept. of Environmental Quality.

Yuan, L. L. 2004. Assigning macroinvertebrate tolerance classifications using generlised additive models. Freshwater Biology **49**:662-677.

Zar, J. H. 1974. Biostatistical Analysis. Prent

Review of: “Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana’s Wadeable Streams and Rivers: Addendum 1”

Peer Review Questions

1. *Approach* - Montana’s approach combines stress-response studies and ecoregional reference distributions to derive numeric nutrient criteria for Montana’s wadeable streams (not really two approaches as identified in review question #1). The approach relies principally on stress-response studies both within the ecoregion and in nearby regions if the reference distributions of TN and TP of the 2 regions are similar. If no stress-response studies were deemed relevant, then MDEQ relied on the reference distribution (e.g., 75th %ile of reference for TP in Absaroka-Gallatin ecoregion 17i). As a further condition on the approach, MDEQ also keeps the N:P ratio in the criteria in a “Redfield range” so that the N:P ratio is unlikely to deviate far from the Redfield ratio, or far from the ecoregional reference if the reference deviates far from Redfield (e.g., Absaroka-Gallatin). The stress-response studies include both experimental nutrient enrichment studies and empirical modeling studies based on monitoring data. Overall, I find the whole approach compelling because it makes effective use of available and relevant information.

   I think the approach would be strengthened by increased use of Montana’s own monitoring data to develop ecoregion-specific empirical models of benthic chl-a response to nutrient enrichment. Such stress-response studies were indeed used if they were available as separate reports or publications, but there is no systematic application of Montana’s data to derive empirical models or confirm the proposed criteria. These models could be used to confirm, refute, or adjust the criteria developed, and would further strengthen the criteria.

   The second part of EPA’s review question, starting with “Please provide documentation on any identified ranges…” is out of line. It is not a reviewer’s task to develop a compendium of alternative methodologies, advantages, disadvantages, data, tools, etc.

2. *River Breaks* – The description of nutrient conditions in the river breaks is plausible, and the river breaks region seems similar (though maybe less extreme) than other badlands ecoregions in the Northwestern Great Plains (including badlands regions of the Dakotas). However, these regions are not familiar to many persons steeped in the Eastern Forest Biome stream paradigms. Accordingly, MDEQ should provide more documenation for the assertions made about the River Breaks. Also, would the same considerations apply to the Little Missouri badlands and the Missouri River Breaks? Evidence could include: published stream studies of badlands-type regions, provided they are similar to the River Breaks; N and P content of soils and geological formations in the River Breaks and similar regions (for example, I have found from EPA’s ecoregion descriptions that the Cretaceous Hell Creek Formation occurs in several of the badlands/breaks ecoregions). Land use/land cover and population density could help show that the breaks and badlands are no different in land use than other Northwestern Great Plains

regions, and perhaps even lower population density and less alteration of land use/land cover than other parts of the Northwestern Great Plains. Early historic descriptions of the regions and their streams are also highly useful, if available.

What is missing from the River Breaks criteria is protection of downstream waters, the large rivers and reservoirs. For example, parts of the region drain into Fort Peck Reservoir. Nutrients in the Breaks streams could contribute to eutrophication of the reservoir. If there are no criteria for the Breaks region, then we could envision the following scenario: In the absence of criteria, the area could become a magnet for large industrial feedlots because no nutrient removal would be required. What happens when hundreds of feedlots drain into the Breaks and on into Ft Peck reservoir? Although criteria may not be required to protect the aquatic life in the streams, they may be required to protect downstream waters.

3. *20% exceedance* – As far as I understand, the 20% exceedance rule means that no more than 20% of single measurements may exceed the nutrient criteria concentrations, or that the site may exceed a criterion up to 20% of the time. Since Montana's proposed criterion is based on single measurements, it is reasonable to expect that some short-term variation above the nutrient criterion concentration will not result in excess chlorophyll. The other alternative is to frame the criteria in terms of an annual average (say, geometric mean) as EPA did for the Florida nutrient criteria. A central tendency measure, geometric or otherwise, also allows for some short-term high concentrations as long as the central tendency is not exceeded. Montana's is basically the same, but based on a percentile of individual measurements.

   The 20% frequency was based on analysis of the Clark Fork River, which shows that for the Clark Fork, the 20% criterion would work well. The problem is that the Clark Fork is a single basin in a restricted set of subecoregions, so we don't have empirical evidence whether 20% would apply to the rest of the state as well. Using Montana's existing monitoring data, I think it may be possible to repeat some of the Clark Fork analysis on other streams throughout the state to confirm or refute the 20% estimate. The Clark Fork data presents another opportunity as well: testing the entire nutrient assessment approach to determine if the actual error rates match with the desired alpha and beta of 0.25 and 0.30. Recommendation: the 20% exceedance rule seems reasonable and has empirical evidence to support it, but would be strengthened by additional analysis from other regions of the state.

4. *125 mg/m^2 chl a* –Benthic chl-a is clearly the most consistent response indicator to nutrients in wadeable streams, as shown by many studies, cited in the MT documents and elsewhere. Benthic macroinvertebrates, while associated with both nutrients and chl-a, have so far proved unsuccessful as a reliable response indicator to nutrient enrichment, as demonstrated by EPA's attempt to develop nutrient criteria for Florida streams. Montana has a rich tradition in monitoring benthic chl-a as well as benthic diatoms, and is making effective use of that tradition for developing nutrient criteria.
   Derivation of the threshold – The threshold was derived from literature values, observations of

streams in the MT ecoregions, an acceptability survey, and a nutrient enrichment study. For example, Welch et al. (1989; cited in MT docs) considered "Nuisance biomass levels" to be in the range 100 – 150 mg/m^2 chl-a. Other values are similar (Biggs 2000: mesotrophy is in the range 60-200 mg/m2; Dodds et al. 2002 [CJFAS 59:865-874]: 125 mg/m^2 is "high end" of chl-*a*). Surveys are context-specific, in that people will identify unacceptable conditions as those that they are not accustomed to seeing. Unaccepatibility thresholds are subject to shifting baselines: if the persons surveyed are accustomed to seeing eutrophic conditions, only hypereutrophy would be identified as unacceptable. Finally, the dose-response study suggested that synoptic reach-average benthic algae in the range 87 – 127 mg/m2 chl a resulted in unacceptable DO at the end of the growing season. These results would suggest that 125 mg/m2 is at or uncomfortably close to a value that could cause fish community degradation due to DO, and for mountain and transitional streams, the chl-a threshold should be lower.
Statewide use - First, it is unclear whether MT plans to use 125 or 150 mg/m^2 as the chl-a standard. Some regions have 125, others 150. As with the nutrient criteria themselves, it may be more appropriate to have chl-a criteria better adjusted to the ecoregions. For example, the expectation for mountain and foothill-transitional ecoregions is that streams are oligotrophic and coldwater, supporting Montana's famous trout fisheries. Given that 125 mg/m2 is in the range of "nuisance", well in "mesotrophy" and has been demonstrated to cause DO problems in Montana, this value is probably too high for mountain and foothill streams. I have no problem with higher values for Plains ecoregions.

5. *Reach-specific criteria* – Two methods for reach-specific criteria are proposed: empirical determination based on pre-defined natural conditions (in this case, dam operations), and ecoregional flow-weighted criteria for streams receiving input from more than a single ecoregion. Both of these approaches appear to be sound.

6. *Tests* – Montana's overall rationale for determining impairment, using both an exact binomial and the t-test, is well thought-out. However, the presentation was a bit confusing; I found I had to jump around between various parts of the 2011 document and its appendixes to understand the approach. The consideration of both significance and power, and the attempt to balance them, is especially encouraging, and shows MDEQ is concerned with both protection of the resource and prevention of unnecessary management. I do have some concerns:

    a. Are the effect size (0.15) and the critical exceedance rate (0.20) really double-counting the same thing? Effect size is a scientific determination that nutrient concentrations within 15% of each other (or within 5% of the criterion) are not meaningfully different in terms of response, so it protects against a statistically significant difference (which may be significant simply due to very large sample size) being declared an impairment when there is actually little chance of impairment for such a small difference. The exceedance rate essentially does the same thing: up to 20% of individual measurements can exceed the criteria, but chl-a will not exceed its criterion value. When both of these are used in the exact binomial, is it testing whether more than 20% of observations exceed the critical nutrient concentration plus 15%? If so, that would be double-counting. The Clark Fork data could be used to test/illustrate this issue empirically. Effect size is typically used in comparisons of central tendency, to protect against scientifically negligible differences being elevated to statistically significant differences simply due to large sample size. It is used most often in equivalence or noninferiority tests. I don't think effect size, as a % of the mean, is appropriate for the exact test, which does not use a mean.

b. Should the effect size be used in the t-test, especially for large or very large sample sizes?

General comments:

In view of several of the questions above, and different ways of calculating status of a streams reach, I was frequently confused whether the document was referring to instantanaeous measures, annual (growing season) maximum, or some measure of central tendency (mean, median, geometric mean, etc.) measure at one time (synoptic) at several sites on a reach, or a "sampling event average". I came to realize that Montana's proposed criteria only make sense in the context of individual measures, i.e., measurements of TN and TP are not to exceed the criterion more than 20%. Similarly, the chl-a criterion of 125 mg/m2 only makes sense as a maximum not-to-be-exceeded. However, it was not clear in the document how exceedance would be calculated. Critically, eventual measures of exceedance should match to the extent possible the way meaningful concentrations were calculated in the considerations to derive the criteria. Recommendation: spell out, with examples, of what is meant by single observations and different central tendencies mentioned in the documents, and which are used for the final criteria and for assessment.

Overall, I found Montana's approach sound and well thought-out. The devil, as always, is in the details: selection of chl-a values, derivation of critical exceedance rates, selection of effect size. Some of the quibbling on these values may never be resolved (including mine), and Montana needs to use best judgment supported by its analysis and other scientific results.