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

Prepared By

Michael Suplee, Ph.D. — Montana Department of Environmental Quality
Vicki Watson, Ph.D. — University of Montana
Arun Varghese and Josh Cleland — ICF International

November 2008
Scientific and Technical Basis of the Numeric Nutrient Criteria for Montana’s Wadeable Streams and Rivers

Prepared By

Michael Suplee, Ph.D.
Montana Department of Environmental Quality
Water Quality Standards Section
1520 E. 6th Ave, Helena, MT 59620-0901

Vicki Watson, Ph.D.
Environmental Studies
University of Montana
32 Campus Drive, Missoula, MT 59812

Arun Varghese and Josh Cleland
ICF International
9300 Lee Highway, Fairfax, VA 22031-1207

November 2008
ACKNOWLEDGMENTS

Many people contributed to the work that has led to this document. We would like to thank all of the people within the Montana Department of Environmental Quality (DEQ) who reviewed an early draft and who provided very useful suggestions for its improvement. We would very much like to thank Walter Dodds (Kansas State University), Michelle Baker (Utah State University), and Jeroen Gerritsen (TetraTech, Inc.) for their careful and insightful reviews of the final draft. We would also like to thank Tina Laidlaw and David Moon (U.S. Environmental Protection Agency, Region VIII) for their reviews. Prior to the preparation of this document, there was a long, sustained effort to gather the essential data upon which it is built. We would therefore like to express our sincere gratitude to the many field crews from the University of Montana and DEQ who, over the years, worked long hours collecting the data which were so critical for the analyses discussed herein. 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.
EXECUTIVE SUMMARY

Beneficial uses are valuable characteristics of a stream or river resource that, directly or indirectly, contribute to human welfare. Examples of beneficial uses include drinking water, fish and aquatic life, and recreation. Beneficial uses are established in law and reflect the societal values embodied in those laws. The intent of water quality criteria, in turn, is to assure a level of water quality that will protect the beneficial uses. Some beneficial uses are more sensitive to impacts (harm) than others; water quality criteria are required by law to protect the most sensitive use from harm. This document presents the science and technical analyses used to develop numeric nutrient criteria for wadeable streams in Montana.

It is well documented that the addition of nitrogen (N) and phosphorus (P) compounds to surface waters leads to a phenomenon referred to as eutrophication. Eutrophication is increased plant and algae growth and decay in a waterbody, and all of the consequential changes to the waterbody and the water quality that occur as a result. N and P criteria are set so that they protect streams from the undesirable aspects of eutrophication. Undesirable aspects of eutrophication include nuisance algae growth and reduced dissolved oxygen levels which impact fish and aquatic life.

Although N and P enrichment causes stream eutrophication, the manner in which eutrophication manifests itself in streams is influenced by other factors. These factors include stream temperatures, flow patterns, light levels, and grazing on algae and plants by fish and aquatic insects. Of these, the most important in Montana appear to be temperature and flow patterns. As such, both temperature and flow patterns were incorporated into the criteria development process; this will be discussed further when the criteria are presented at the end of this summary.

Montana already has several water quality standards that address undesirable aspects of eutrophication. (Water quality standards are, essentially, criteria that have been adopted into law.) However, the existing standards are either narrative (they describe a water quality condition that should be maintained, but provide no specifics, therefore they are open to varied interpretations), or they are numeric but address only effect variables (e.g., low dissolved oxygen levels). Thus, one is still required to determine the root cause of the effects and that root cause is commonly nutrient enrichment. Numeric nutrient criteria will improve upon the existing standards because they address the causes of eutrophication directly.

N and P concentrations in wadeable streams vary naturally in accordance with regional geology, soils, climate, and vegetation. To address this regional variation, DEQ tested three candidate mapping systems. The intended purpose of the mapping system was to assure that appropriate nutrient criteria are applied to different regions of the state given the natural spatial variation in nutrient concentrations. Omernik ecoregions, Strahler stream order, and underlying geology (lithology) were each evaluated to see which one was the best at maximizing the difference in stream nutrient concentrations between zones and minimizing the difference within zones. Among the three mapping systems, ecoregions were found to be superior to the others and are recommended as the system upon which nutrient criteria zones should be based.
In some parts of the state, mainly in the west, the most sensitive beneficial use is recreation. A public opinion study carried out by DEQ shows that the public majority in Montana does not want to see excessive bottom-attached algae growth in the gravel-bottomed, clear running, trout-fishery streams common in western Montana. For these types of streams, the nutrient criteria have been set to prevent nuisance algal levels (as defined by the public perception study) from developing and will, therefore, protect the recreation use. The criteria will also protect the fishery, which typically comprises fish such as trout, char, and whitefish, from the negative effects of excessive nutrient enrichment (e.g., low dissolved oxygen concentrations). The criteria will also better protect the agricultural use by reducing elevated algae levels that clog irrigation systems.

In the eastern part of the state, low gradient prairie streams are common. Wadeable prairie streams in Montana often become intermittent, commonly have mud bottoms, are turbid, frequently have substantial macrophyte populations, usually have filamentous algae but sometimes have only phytoplankton algae, and support catfish, walleye, chubs, bass, and other warm water fishes. Because prairie streams are fundamentally different in many ways from western Montana trout streams, the results from the algae public perception survey should probably not be directly applied to them. Prairie streams nevertheless have important and sensitive beneficial uses that need protection, like the diverse species of fish mentioned above. For these types of streams, the nutrient criteria have been set so that they will maintain dissolved oxygen concentrations that protect regional fish and aquatic life. The most sensitive use in prairie streams is therefore considered to be fish and aquatic life.

Fundamentally, the nutrient criteria are based on scientific stressor-response studies in which harm to a sensitive beneficial use is shown. All applicable stressor-response studies (N or P as stressor, beneficial use impact as response) that could be located were reviewed. This included regional studies as well as studies from other parts of the country and world. Some of these studies were carried out by DEQ.

We then compared the nutrient concentrations indicated by the stressor-response studies to regionally-applicable reference stream nutrient data. This analysis showed that there is a consistent relationship between nutrient concentrations that harm uses (as determined in the stressor-response studies) and nutrient concentrations observed in reference sites; namely, the most elevated nutrient concentrations observed in reference sites (e.g., those at the 87th percentile of reference) are equivalent to harm-to-use concentrations. It is not surprising that reference sites have some nutrient samples whose concentrations are higher than the harm-to-use threshold identified in stressor-response studies. In any population there are always low and high values that differ considerably from the population’s central tendency; the important point is that nutrient concentrations in reference sites that are greater than the harm-to-use threshold occur infrequently, e.g. due to an atypical high-flow event in summer. It is when the harm-to-use concentrations occur commonly in a stream (e.g., 50% of the time) that eutrophication problems occur.

Owing to the collective relationship we observed between stressor-response studies and the their corresponding reference data, we recommend using nutrient concentrations linked to specified upper percentiles of the reference data (e.g., the 90th of reference) as criteria. This approach has
the advantage that it helps overcome statistical uncertainties in any given stressor-response study, and makes certain that natural, regional effects on background nutrient concentrations are reflected in the criteria so that the criteria will not be overly stringent or insufficiently protective.

Table E1 below shows the recommended numeric criteria for different ecoregions. Both total P and total N are recommended, as co-limitation by both nutrients is common in rivers and streams. Nitrate + nitrite (NO$_{2-3}$) is also suggested because total N criteria — in the absence of NO$_{2-3}$ criteria — may not achieve the water quality goals anticipated (nitrate is particularly important in prairie streams). Analysis shows that the total P criteria should generally maintain soluble P at appropriate levels.

The criteria only apply seasonally. This is because low temperatures in winter and high flow events during spring runoff tend to mute the local effects of eutrophication (plant growth slows dramatically in winter, and spring high-flow events prevent nuisance algal mats from developing). Therefore, the criteria have been set for the time period when eutrophication problems are most likely to occur (i.e., summer). Note also that we have included benthic (i.e., bottom-attached) algae criteria for the mountainous ecoregions (Northern, Canadian, and Middle Rockies, and the Idaho Batholith). The algae levels shown are based on DEQ’s nuisance algae public-perception survey. In these mountainous ecoregions, the algae criteria should be adopted along with the nutrient criteria to assure protection of the beneficial uses. In the eastern prairie streams (Northwestern Glaciated Plains, Northwestern Great Plains, and Wyoming Basin), nutrient criteria are provided but benthic algae levels are not. As noted earlier, prairie stream nutrient criteria are intended to maintain dissolved oxygen levels already in state law and are not based on the algae public-perception survey. As for all water quality criteria, the numeric nutrient criteria will undergo periodic revision and update as more stressor-response studies are completed and more reference data are collected.

<table>
<thead>
<tr>
<th>Level III Ecoregion</th>
<th>Period When Criteria Apply</th>
<th>Total P (mg/L)</th>
<th>Total N (mg/L)</th>
<th>NO$_{2-3}$ (mg/L)</th>
<th>Benthic Algae Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Rockies</td>
<td>July 1 - Sept. 30</td>
<td>0.012</td>
<td>0.233</td>
<td>0.081</td>
<td>150 mg Chl $a/m^2$ (36 g AFDW/m$^2$)</td>
</tr>
<tr>
<td>Canadian Rockies</td>
<td>July 1 - Sept. 30</td>
<td>0.006</td>
<td>0.209</td>
<td>0.020</td>
<td>150 mg Chl $a/m^2$ (36 g AFDW/m$^2$)</td>
</tr>
<tr>
<td>Middle Rockies</td>
<td>July 1 - Sept. 30</td>
<td>0.048</td>
<td>0.320</td>
<td>0.100</td>
<td>150 mg Chl $a/m^2$ (36 g AFDW/m$^2$)</td>
</tr>
<tr>
<td>Idaho Batholith</td>
<td>July 1 - Sept. 30</td>
<td>0.011</td>
<td>0.130</td>
<td>0.049</td>
<td>150 mg Chl $a/m^2$ (36 g AFDW/m$^2$)</td>
</tr>
<tr>
<td>Northwestern Glaciated Plains</td>
<td>June 16 - Sept. 30</td>
<td>0.123</td>
<td>1.311</td>
<td>0.020</td>
<td>n/a</td>
</tr>
<tr>
<td>Northwestern Great Plains, Wyoming Basin</td>
<td>July 1 - Sept. 30</td>
<td>0.124</td>
<td>1.358</td>
<td>0.076</td>
<td>n/a</td>
</tr>
</tbody>
</table>
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Section 1.0 Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Scope of the Criteria</td>
<td>1</td>
</tr>
<tr>
<td>1.3 Nutrient Criteria are Different than Other Water Quality Criteria</td>
<td>2</td>
</tr>
<tr>
<td>1.4 Beneficial Uses and Criteria</td>
<td>3</td>
</tr>
<tr>
<td><strong>Section 2.0 The Science of Stream Eutrophication</strong></td>
<td>5</td>
</tr>
<tr>
<td>2.1 Stream Eutrophication, Nitrogen, and Phosphorus</td>
<td>5</td>
</tr>
<tr>
<td>2.2 How Eutrophication Manifests in Streams</td>
<td>5</td>
</tr>
<tr>
<td>2.3 The Influence of Other Stream Environmental Factors on Eutrophication</td>
<td>7</td>
</tr>
<tr>
<td><strong>Section 3.0 Existing vs. Proposed Approach to Controlling Stream Eutrophication</strong></td>
<td>11</td>
</tr>
<tr>
<td>3.1 The Push for Numeric Nutrient Criteria at the National Level</td>
<td>11</td>
</tr>
<tr>
<td>3.2 How Eutrophication has been Addressed in Montana Up Till Now, and How Numeric Nutrient Criteria Will Improve the State’s Existing Water Quality Standards</td>
<td>11</td>
</tr>
<tr>
<td>3.3 Recommended Nutrients, and Determining Compliance with the Criteria</td>
<td>12</td>
</tr>
<tr>
<td>3.4 Difficulties with Numeric Nutrient Criteria</td>
<td>13</td>
</tr>
<tr>
<td><strong>Section 4.0 How DEQ Selected a Geographic Stratification System to Apply Different Nutrient Criteria in Different Places</strong></td>
<td>15</td>
</tr>
<tr>
<td>4.1 Purpose of Developing a Geographic Stratification System</td>
<td>15</td>
</tr>
<tr>
<td>4.2 Conceptual Approaches for Developing a Geographic Stratification System...</td>
<td>16</td>
</tr>
<tr>
<td>4.2.1 The Random Delineation Approach to Establishing Regional Nutrient Zones</td>
<td>16</td>
</tr>
<tr>
<td>4.2.2 The Factor Analysis Approach to Establishing Regional Nutrient Zones</td>
<td>17</td>
</tr>
<tr>
<td>4.2.3 The Combined Approach to Establishing Regional Nutrient Zones</td>
<td>17</td>
</tr>
<tr>
<td>4.2.4 DEQ’s Preferred Approach</td>
<td>18</td>
</tr>
<tr>
<td>4.3 Extant Classification Systems</td>
<td>18</td>
</tr>
<tr>
<td>4.4 Historical Database of Montana Stream Monitoring Data</td>
<td>20</td>
</tr>
<tr>
<td>4.5 Statistical Methods used to Test the Geographic Nutrient Zones</td>
<td>21</td>
</tr>
<tr>
<td>4.5.1 Non-Parametric Tests for Determining Differences Between Stratified Populations</td>
<td>21</td>
</tr>
<tr>
<td>4.5.2 Parametric Tests for Determining Differences Between Stratified Populations</td>
<td>22</td>
</tr>
<tr>
<td>4.5.3 Computation of Measures of Variance for Alternative Stratification Methods</td>
<td>23</td>
</tr>
<tr>
<td>4.6 Other Issues Affecting the Choice of Nutrient Zones</td>
<td>24</td>
</tr>
<tr>
<td>4.6.1 The Ideal Number of Regional Nutrient Zones</td>
<td>24</td>
</tr>
<tr>
<td>4.6.2 Nutrient-Specific Zones</td>
<td>25</td>
</tr>
<tr>
<td>4.6.3 Seasonal Considerations</td>
<td>25</td>
</tr>
<tr>
<td>4.7 Results of the Empirical Analysis</td>
<td>26</td>
</tr>
<tr>
<td>4.8 Conclusions about the Geographic Stratification Systems</td>
<td>27</td>
</tr>
<tr>
<td><strong>Section 5.0 Identifying Eutrophication Impacts on Sensitive Beneficial Uses</strong></td>
<td>28</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Diel Dissolved Oxygen (DO) Patterns in a Montana Stream, 2003. The DO concentrations at 8 and 4 mg/L are intended to protect juvenile and adult fish, respectively.</td>
<td>4</td>
</tr>
<tr>
<td>Figure 2.1</td>
<td>Example of Heavy Cladophora Growth in a Wadeable Montana River (Aug., 2004).</td>
<td>6</td>
</tr>
<tr>
<td>Figure 2.2</td>
<td>Number of Salmonid Fish Per Unit Stream Length in Wisconsin Streams. Modified from Figure 3 in Wang et al. (2007) to show greater resolution at lower total P concentrations. Presented with permission of the authors.</td>
<td>7</td>
</tr>
<tr>
<td>Figure 2.3</td>
<td>Influence of Light Attenuation on Productivity in a Montana River. Unusually high turbidity dampened DO changes over a three day period. During the 4th day, turbidity dropped to more typical background levels and the magnitude of the diel DO oscillations (i.e., productivity) increased markedly.</td>
<td>9</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Omernik Level III Ecoregions in Montana. Level IV ecoregions are shown as outlined areas within each level III ecoregion. Reference sites in Montana through 2007 are identified with black dots.</td>
<td>20</td>
</tr>
<tr>
<td>Figure 6.1</td>
<td>Distribution of Strahler Stream Orders for Montana Reference Sites.</td>
<td>33</td>
</tr>
<tr>
<td>Figure 6.2</td>
<td>Conceptual Diagram Showing a Nutrient Concentration Histogram for Reference Sites. The figure shows where along the x-axis, relative to the histogram, nutrient concentrations likely to harm beneficial water uses would be expected to be found.</td>
<td>35</td>
</tr>
<tr>
<td>Figure 7.1</td>
<td>Willow Creek. A reference stream site in the level III ecoregion Middle Rockies.</td>
<td>43</td>
</tr>
<tr>
<td>Figure 7.2</td>
<td>Rock Creek, a Reference Prairie Stream Site in the Level III Ecoregion Northwestern Glaciated Plains</td>
<td>45</td>
</tr>
<tr>
<td>Figure 7.3</td>
<td>Rock Creek, a Reference Stream Site in the Transitional Level IV.</td>
<td>46</td>
</tr>
<tr>
<td>Figure 7.4</td>
<td>Sweet Grass Creek, a Reference Stream Site in the Transitional Level IV Ecoregion Non-calcareous Foothill Grassland (Part of the Northwestern Great Plains).</td>
<td>46</td>
</tr>
<tr>
<td>Figure 7.5</td>
<td>Mountain-to-Prairie Transitional Level IV Ecoregions. The level IV transitional ecoregions are shown in dark brown. Reference sites are white dots.</td>
<td>47</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1.1</td>
<td>Large River Segments in Montana. This Table May Be Subject to Further Refinement.</td>
<td>2</td>
</tr>
<tr>
<td>Table 4.1</td>
<td>Start and Ending Dates for Three Seasons (Winter, Runoff and Growing), by Level III Ecoregion.</td>
<td>26</td>
</tr>
<tr>
<td>Table 6.1</td>
<td>Studies Addressing Nutrients and Eutrophication that Were Useful for Developing Montana’s Numeric Criteria.</td>
<td>31</td>
</tr>
<tr>
<td>Table 6.2</td>
<td>Nutrient Sampling Indices for Montana Reference Sites, Before &amp; After 2007 Sampling. Data are Presented by Level III Ecoregion.</td>
<td>34</td>
</tr>
<tr>
<td>Table 6.3</td>
<td>Descriptive Statistics for Nutrients from Montana Reference Sites, 2008 Database. Data are Presented by Level III Ecoregion.</td>
<td>36</td>
</tr>
<tr>
<td>Table 6.4</td>
<td>Regional Stresor-Response Study Nutrient Concentrations and their Corresponding Percentile Values in Corresponding Reference-Site Nutrient Frequency Distributions.</td>
<td>37</td>
</tr>
<tr>
<td>Table 6.5</td>
<td>Nutrient Concentrations from Studies Carried out in Northern Temperate Rivers &amp; Streams Compared to Nutrient Concentrations at the 90th Percentile of Reference Sites from Montana’s Proportion of the Middle Rockies Ecoregion.</td>
<td>40</td>
</tr>
<tr>
<td>Table 7.1</td>
<td>Example Criteria for Different Ecoregions in Montana. The Numeric Nutrient Criteria Values May Change Slightly Due to Ongoing Data Collection in Reference Sites.</td>
<td>42</td>
</tr>
</tbody>
</table>
GLOSSARY OF TERMS AND MEASUREMENT UNITS

Glossary of Terms

ABFW: Ash free dry weight. The weight (mass) of a material obtained by first drying the material at 105 °C and then heating the material to 500 °C for 1 hour.

Algae: Aquatic organisms that photosynthesize but lack a vascular system. Some are microscopic, others very large.

Aquatic Macroinvertebrate: An organism found in waterbodies, frequently associated with stream bottoms, not having a spinal column and which is visible with the naked eye.


Baseflow: The portion of streamflow that comes from groundwater and not runoff.

Beneficial Use: A valuable characteristic of a stream or river resource that, directly or indirectly, contributes to human welfare.

Benthic: On or associated with the sediments or bottom of a body of water.

Biomass: The total mass or amount of living or dead organisms in a particular area or volume.

BOD (Biochemical Oxygen Demand): A measure of, as well as a procedure for determining, how fast oxygen is used up in water. BOD is usually measured as the rate of oxygen uptake by microorganisms in a water sample, at 20°C, in the dark, over a five day period.


Chlorophyll a: The major green pigment found in the chloroplasts of plants and algae.

Density: Quantity of a number per unit area, volume, or mass.

Diatom: Any one of a number of microscopic algae, which can live as single cells or in colonies, that are enclosed within two box-like parts or valves (called frustules) made of silica that fit together like the halves of a Petri dish.

Diel: Involving a 24-hour period that usually includes a day and the adjoining night.

Ephemeral Stream: A stream or stream segment which flows only in direct response to precipitation in the immediate watershed or in response to the melting of a cover of snow and ice and whose channel bottom is always above the local water table.

Eutrophication: The process of enrichment of a waterbody by nutrients, usually nitrogen and phosphorus containing compounds, and the resulting increase in primary productivity (algal growth, etc.).
and plant growth and decay). Some definitions include organic enrichment of a waterbody as part of eutrophication.

**Fixation (Nitrogen Fixation):** The process by which nitrogen is taken from its relatively inert gas form in the atmosphere (N$_2$) and converted into nitrogen compounds such as nitrate and ammonia.

**Geospatial:** Pertaining to the geographic location and characteristics of natural or constructed features and boundaries on, above, or below the earth's surface; especially referring to data that is geographic and spatial in nature.

**Intermittent Stream:** A stream or stream segment that is below the local water table for at least some part of the year, and obtains its flow from both surface run-off and ground water discharge.

**Macrophyte:** Macroscopic aquatic vascular plants capable of achieving their generative cycles with all or most of the vegetative parts submerged or supported by the water.

**Mainstem:** The principal river within a given drainage basin, in the case where a number of tributaries discharge into a larger watercourse.

**MCA:** Montana Code Annotated.

**Metric:** A characteristic of a biological assemblage (e.g., fishes, algae) that changes in some predictable way with increased human influence.

**Narrative Water Quality Criteria:** Statements codified in state law that describe, in a concise way, a water quality condition that must be maintained in order to protect beneficial uses.

**Nonpoint Source:** The source of pollutants which originates from diffuse runoff, seepage, drainage, or infiltration.

**Numeric Water Quality Criteria:** Quantified expressions of water quality, in state law, intended to protect a designated beneficial use or uses.

**Organic Enrichment:** From a water pollution perspective, the addition of decomposable plant or animal material, or their wastes, to a waterbody.

**Perennial Stream:** A stream or stream segment that has flowing water year-round except during extreme drought.

**Periphyton:** The microscopic flora and fauna that grow or are associated with the bottom of a body of water, and includes microscopic algae, bacteria, and fungi.

**Phytoplankton:** Free living, generally microscopic algae commonly found floating or drifting in waterbodies such as the ocean, lakes and streams.
**Point Source:** A discernable, confined, and discrete conveyance, including but not limited to any pipe, ditch, channel, tunnel, conduit, well, discrete fissure, container, rolling stock, or vessel or other floating craft, from which pollutants are or may be discharged.

**Primary Productivity:** The production of organic compounds from carbon dioxide, principally through the process of photosynthesis.

**Redfield Ratio:** The molecular ratio of carbon, nitrogen and phosphorus in phytoplankton. The molar ratio is 106:16:1 (C:N:P), or 47:7:1 by weight. The term is named after the American oceanographer Alfred C. Redfield.

**Salmonids:** Ray-finned fish, whose members include salmon, trout, chars, freshwater whitefishes and graylings.

**Saturation:** A state in which the gas concentration (e.g., oxygen) in a waterbody is in equilibrium with the local partial pressure of that gas in the atmosphere.

**Standards (Water Quality Standards):** In a water quality regulatory context, a term applicable to state waters referring collectively to their designated beneficial uses, criteria, and the non-degradation policy, all in Montana law.

**Strahler Order:** A simple hydrology algorithm used to define stream size based on a hierarchy of tributaries. Streams at the top of the watershed are labeled 1. When two order-1 streams join, they create an order-2 stream. When two order-2 streams join, they create an order-3 stream, and so on. If a stream of lower order (e.g., order-2) joins a stream of higher order (e.g., order-3), the order number of the latter does not change.

**Wadeable:** A stream whose Strahler order is first through (at most) sixth (1:100,000 map scale) in which most of the wetted channel is wadeable by a person during baseflow conditions.

### Measurement Units

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Unit Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cm</td>
<td>centimeter</td>
</tr>
<tr>
<td>ft</td>
<td>feet</td>
</tr>
<tr>
<td>in</td>
<td>inch</td>
</tr>
<tr>
<td>hr</td>
<td>hour</td>
</tr>
<tr>
<td>L</td>
<td>liter</td>
</tr>
<tr>
<td>m</td>
<td>meter</td>
</tr>
<tr>
<td>m²</td>
<td>square meter</td>
</tr>
<tr>
<td>m³</td>
<td>cubic meter</td>
</tr>
<tr>
<td>mg</td>
<td>milligram</td>
</tr>
<tr>
<td>mm</td>
<td>millimeter</td>
</tr>
<tr>
<td>NTU</td>
<td>nephelometric turbidity units</td>
</tr>
<tr>
<td>µg</td>
<td>microgram</td>
</tr>
</tbody>
</table>
1.0 Introduction

Section 1.0
Introduction

1.1 Background

The Montana Department of Environmental Quality (DEQ) has developed numeric water quality criteria intended to control excessive nutrient pollution in Montana’s wadeable rivers and streams. This document describes why numeric nutrient criteria are needed and DEQ’s technical basis for developing them. The criteria are the culmination of eight years of work by DEQ. Because of the potential regulatory implications of nutrient criteria, DEQ has taken a measured, cautious approach in developing them and has strived to base them on the best available science and data. A great deal of scientific understanding about the role of nutrients in stream ecology already existed at the onset of this work, in 2000. Since that time research carried out regionally, nationally, and internationally has only further increased the scientific knowledge base. Some of this scientific work has been carried out in Montana. Of equal or perhaps greater importance has been the increased clarification of how, and at what point, surface water resources are harmed by excess nutrients. Among water quality managers, the term “beneficial use” is often heard, and will be repeated throughout this document. Beneficial uses are valuable characteristics of a stream or river resource that, directly or indirectly, contribute to human welfare. Beneficial uses are known by other names (e.g. instream values, valued ecological attributes), but one basic truth remains the same; determining the threshold that defines when an impact to a beneficial use has occurred requires both scientific understanding and value judgment. Thus, clarity about how, when, and why beneficial uses become harmed by nutrients is at the heart of setting numeric nutrient criteria. In writing this document, it was our intent that the reader gains a basic understanding of the ecological role of nutrients in streams, how beneficial uses become harmed by nutrients, and how the criteria are expected to prevent the latter from occurring.

1.2 Scope of the Criteria

The criteria discussed in this document apply specifically to wadeable streams. A wadeable stream is generally defined here as a stream whose Strahler order is first through (at most) sixth (1:100,000 map scale) in which most of the wetted channel is wadeable by a person during baseflow conditions. This includes perennial streams that run all year and streams that become intermittent, i.e., a disconnected series of pools. The criteria do not apply to ephemeral channels, defined in Montana as a stream or stream segment which flows only in direct response to precipitation in the immediate watershed or in response to the melting of a cover of snow and ice and whose channel bottom is always above the local water table (ARM 17.30.602[12]). Nor do the criteria apply to large rivers, which are generally 7th order or larger and in which most of the wetted channel is unwadeable during baseflow; examples include the Yellowstone River and the Missouri River (Table 1.1). Finally, the criteria do not apply to lakes or wetlands. Separate efforts are currently underway to develop criteria for large rivers and lakes, and will be addressed in a future document similar to this one.
1.0 Introduction

Table 1.1 Large River Segments in Montana. This Table May Be Subject to Further Refinement.

<table>
<thead>
<tr>
<th>River Name</th>
<th>Segment Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clark Fork River</td>
<td>From Rock Creek confluence (46.726, -113.683) to Idaho border*</td>
</tr>
<tr>
<td>Flathead River</td>
<td>From the Flathead Indian Reservation boundary to the mouth</td>
</tr>
<tr>
<td>Kootenai River</td>
<td>From Libby Dam to Idaho border</td>
</tr>
<tr>
<td>Marias River</td>
<td>From Tiber Dam to the mouth</td>
</tr>
<tr>
<td>Musselshell River</td>
<td>From Flatwillow Creek confluence (46.928, -107.930) to Fork Peck Reservoir</td>
</tr>
<tr>
<td>Milk River</td>
<td>From Fresno Dam to the mouth</td>
</tr>
<tr>
<td>Missouri River</td>
<td>Entire length in Montana</td>
</tr>
<tr>
<td>Bighorn River</td>
<td>From Yellowtail Dam to Crow Indian Reservation boundary near St. Xavier, MT, and from the Crow Indian Reservation boundary near Hardin, MT, to the mouth</td>
</tr>
<tr>
<td>Tongue River</td>
<td>From Hanging Woman Creek confluence (45.321, -106.522) to the mouth</td>
</tr>
<tr>
<td>Powder River</td>
<td>From Wyoming border to the mouth</td>
</tr>
</tbody>
</table>

*Numeric nutrient and algal biomass standards already are in place from the Rock Cr confluence downstream to the Flathead River confluence (ARM 17.30.631)

1.3 Nutrient Criteria are Different than Other Water Quality Criteria

Nitrogen and phosphorus are essential nutrients for plants and animals. Without them, organisms would not be able to build the proteins and nucleic acids of their cellular structures or carry out the basic oxidation and reduction reactions that power each of their cells. Because nitrogen and phosphorus are necessary for all organisms, the effects these nutrients have in the aquatic environment are inherently different from the effects of toxic substances. Many substances (e.g., lead and mercury) are toxic to people or aquatic organisms in the tiniest of concentrations and, traditionally, water quality criteria for these types of elements and compounds have been derived by toxicologists using laboratory studies. Some nutrients (e.g., certain nitrogen compounds) can be, at fairly high concentrations, toxic to people and aquatic organisms as well. But criteria have already been established for those effects. The nitrogen and phosphorus criteria presented in this document are intended to control the undesirable aspects of an environmental effect referred to as eutrophication (or sometimes “cultural” eutrophication).

Eutrophication is the enrichment of a waterbody by (typically) nitrogen and phosphorus, leading to increased plant and algae growth and decay, and all the consequential changes to the waterbody and the water quality that occur as a result of this enrichment. Enrichment of waterbodies by nutrients is not in and of itself negative. Many waterbodies are purposefully enriched in order to enhance their overall productivity; fertilization of commercial fish ponds to increase growth of certain fish species is such an example. Enrichment becomes detrimental — and therefore begins to fall within the realm of water quality criteria setting — when the effects manifested in a waterbody are undesirable relative to the uses the waterbody is intended for. Nitrogen and phosphorus concentrations in waterbodies have the interesting quality that as they increase they may initially enhance certain waterbody characteristics (e.g., cause fish to grow

---

\(^a\) Water quality criteria are numeric or narrative expressions of water quality that, if achieved, assure that important characteristics of a water resource are not damaged.
larger) but then, at higher concentrations, lead to conditions that harm the value of the waterbody (e.g., result in low dissolved oxygen that impairs the fishery). Thus, setting numeric nutrient criteria requires an understanding of how eutrophication progresses with increasing nutrient concentrations, and at what point the detrimental effects begin to occur. This document will discuss each of these issues in detail.

1.4 Beneficial Uses and Criteria

Identifying when water quality changes begin to have detrimental effects on stream resources is key to setting appropriate water quality criteria. But even before a detrimental effect or criterion can be considered, some frame of reference is needed to help set the expectations. This frame of reference is established via the stream’s beneficial uses. Beneficial uses are the valuable characteristics of a stream or river resource that, directly or indirectly, contribute to human welfare. They are established in law and reflect the societal values embodied in those laws. Some typical examples of beneficial uses in streams are public water supply, fish and aquatic life, agricultural use, and recreational use. Montana has had this very list of beneficial uses in law since 1955. Montana records its beneficial-use definitions in its Administrative Rules (ARMs) at 17.30.601 et seq. Beneficial uses became more uniformly defined nationally when the U.S. Congress amended the Federal Water Pollution Control Act in 1972, leading to what is commonly called the Clean Water Act. The U.S. Clean Water Act clearly states that water quality should be protected so that fish propagation and recreation (i.e., fishable and swimmable) would be achieved in all waterbodies, where possible.

Once the beneficial uses of waterbodies are established, the framework is in place to develop mechanisms to protect the beneficial uses from harm. This leads directly back to water quality criteria. As noted earlier, criteria are numeric or narrative expressions of water quality that, if achieved, assure that the uses are not harmed by human action. As might be expected, some uses are more sensitive to impacts than others. For example, both the fish and aquatic life use and the industrial use are beneficial water uses. But fish are much more sensitive to many water quality changes (e.g., increases in copper concentration) than most industrial processes would be to the same changes. Thus, the fishery use is — in this example — the most sensitive use. In Montana all uses are valued the same; i.e., all uses are treated equally. Water quality criteria in Montana are set to protect the most sensitive use, with the understanding that the less sensitive uses will be protected automatically; this approach is also required under Federal law (40 CFR 131.11). This method was used in developing the numeric nutrient criteria presented in this document.

As states, tribes, and the U.S. Environmental Protection Agency (EPA) have developed water quality criteria to protect beneficial uses, they have used a variety of techniques to identify when harm occurs to the use. These techniques are a function of the use in question and the specific technical approaches required. But these techniques all have in common the integration of value judgments and science. Beneficial uses and their associated impact (or harm) thresholds reflect societal values codified in law, while the scientific method provides understanding as to how particular water quality parameters affect streams and quantifies when those parameters harm the use. Figure 1.1 provides an example of harm to a beneficial use (salmonid fishes). Large day-to-night dissolved oxygen (DO) changes measured in this Montana stream were linked to human pollution causes; note how they go over 14 mg/L in the day and drop below 4 mg/L at night. It is...
well established scientifically that low DO concentrations harm fish\textsuperscript{6,7}; for example, weight gain and food conversion (amount consumed to amount of weight gained) of young salmonid fishes drops rapidly when DO is below 4 mg/L\textsuperscript{6}. For the stream in Figure 1.1, DO levels of 8 mg/L for juvenile fish and 4 mg/L for adult fish are instantaneous minima in Montana law\textsuperscript{8} intended to assure proper egg development\textsuperscript{b} and normal growth and activity for completing all life stages of fish like rainbow and brook trout. The DO concentrations in the stream in Figure 1.1 fall below these criteria and are, therefore, harming the beneficial use.

\textbf{Figure 1.1}  
**Diel Dissolved Oxygen (DO) Patterns in a Montana Stream, 2003.** The DO concentrations at 8 and 4 mg/L are intended to protect juvenile and adult fish, respectively.

\textsuperscript{b} Montana water quality standards require 8 mg DO/L in the water in order to achieve 5 mg DO/L in the inter-gravel region where embryonic, larval, and juvenile fish are commonly found.
Section 2.0
The Science of Stream Eutrophication

2.1 Stream Eutrophication, Nitrogen, and Phosphorus

Eutrophication (nutrient enrichment) management efforts focus on nitrogen (N) and phosphorus (P) concentrations and not on other nutrients (e.g., trace minerals, vitamins, etc.) for good reason. After much scientific research concerning which major nutrients most often limit or control primary productivity in freshwaters, there is general consensus that it is typically P and or N. Major nutrients required by organisms are carbon, oxygen, hydrogen, nitrogen, and phosphorus and, at various times, limnologists (scientists that study fresh waters) have considered most of these elements as candidates for controlling productivity in freshwaters. In the 1960s there was considerable debate about the relative importance to primary productivity of inorganic carbon vs. phosphorus, but Schindler’s influential work on whole-lake fertilization showed that phosphorus limited lake productivity, and carbon did not. After Schindler’s lake work in the 1970s, there was a general emphasis on P as the sole nutrient limiting productivity and controlling eutrophication in freshwater systems. However, N has in more recent decades been found to be of equal importance in rivers and streams, where N and P co-limitation appears to be common. This has been further confirmed by whole-river and whole-stream fertilization experiments using N and P that demonstrate that production in flowing waters is strongly controlled by N and P. Plant physiologists have identified a number of required micronutrients that could potentially limit freshwater productivity, but these substances are needed in minute quantities and field studies show they do not limit stream production; thus, it is very unlikely that they generally limit freshwater productivity. Site-specific exceptions to the previous statement may exist, e.g. in some granitic alpine areas, but these are not widespread enough to warrant consideration for water quality management purposes.

Nitrogen and P enter streams by a variety of routes. Point sources (e.g., waste water treatment plants) with an end-of-pipe discharge are probably the most conspicuous. But they are not the only sources and, in certain situations, are not even the largest contributors. Since the end of WW II humans have dramatically increased the fixation of N (e.g., via the Haber-Bosh industrial process which converts unreactive atmospheric N to ammonia salts) to the point that human sources of fixed N now exceed natural sources on a global scale. Most of this fixed N is used for growing food crops and, as a result, large amounts of N have entered streams via point and nonpoint sources. Agricultural sources are a major sources of N and P to streams, and are often a mix of organic and inorganic forms, while other nonpoint nutrient sources include soil and stream bank erosion, urban runoff and sprawl, land clearing and conversion, and loss of wetlands and the subsequent oxidation of their organic soils.

2.2 How Eutrophication Manifests in Streams

Eutrophication has a number of effects in flowing waters. One very typical change is the increased dominance by benthic (i.e., bottom attached) filamentous algae in temperate streams. The green algae Cladophora spp. in particular seems to benefit from increased nutrient enrichment. Cladophora (Figure 2.1) probably played a minor role in aquatic communities...
The Science of Stream Eutrophication

Figure 2.1 Example of Heavy Cladophora Growth in a Wadeable Montana River (Aug., 2004).

before widespread cultural eutrophication occurred \(^{20,21}\), but dense growths of the alga have now become common in nutrient-enriched temperate streams worldwide, including in Montana \(^{20,22-24}\). Another common effect of eutrophication in streams and rivers is the increased magnitude of daily DO and pH oscillations due to the elevated productivity of phytoplankton, benthic algae, or both \(^{25}\). Aquatic insect (aquatic macroinvertebrate) populations often shift in response to increasing nutrient enrichment and there is a large scientific literature devoted to the relationships between macroinvertebrates and water quality. Very generally, sensitive macroinvertebrates, including mayflies (Ephemeroptera), stoneflies (Plecoptera), and caddisflies (Trichoptera), tend to be found in clean water having low nutrient concentrations and DO near saturation (i.e., without extreme daily oscillations). At the other end of the spectrum, many midge species (chironomids) are tolerant of heavy eutrophication and the associated conditions (e.g., low nighttime DO) \(^{21,26-28}\). Downstream of efficiently-treated wastewater discharges where the effluent contains mainly inorganic N and P (i.e., little organic matter), overall macroinvertebrate density and biomass increase, but the density of pollution sensitive species diminishes \(^{29}\).

Fish populations are also affected by eutrophication. It has been shown that increasing N and P concentrations in streams can result in increased growth of fish, including salmonid fishes (e.g., trout, char) \(^{13,14,30,31}\). Phosphorus is shown to have a threshold effect on salmonid fish.
populations, whereby the number of fish found per 100 m of stream length initially increases with increasing P, but then drops after P reaches a certain threshold concentration\(^5\) (Figure 2.2). Ultimately, if eutrophication becomes too severe, fish kills can occur due to low nighttime DO\(^3\). Studies also show that stream eutrophication (quantified as increasing benthic algae density) leads to greater accumulation of organochlorine pollutants (e.g., PCBs) in localized trout populations. This occurs because increased primary productivity in rivers slows downstream transport of the PCBs, allowing fish more time to ingest them\(^33,34\).

![Figure 2.2](image)

**Figure 2.2** Number of Salmonid Fish Per Unit Stream Length in Wisconsin Streams. Modified from Figure 3 in Wang et al. (2007) to show greater resolution at lower total P concentrations. Presented with permission of the authors.

### 2.3 The Influence of Other Stream Environmental Factors on Eutrophication

Non-nutrient environmental factors influence the way eutrophication manifests itself in streams. At a very coarse scale, climate, source of flow, and geology set the stage for the types of streams that develop in a region\(^35\). Within any particular stream, factors external to nutrients that influence primary productivity include light, temperature, water velocity and high-flow events, grazing by fish and macroinvertebrates, and inter-specific competition among the plants\(^36\). Those we considered among the most important to Montana streams are detailed below.

When sufficient light reduction occurs, eutrophication effects are muted, at least locally, because aquatic plants driving the productivity become light limited\(^37-39\). However, light must be reduced by 60% or more from ambient levels in order to reduce algae proliferations\(^2,39\). Many smaller streams in western Montana have riparian canopies which substantially diminish light levels at the stream surface, but it is also true that un-canopied areas are commonly interspersed along their lengths as they flow through open meadows and pastures. Light attenuation is provided by riparian canopy, but it can also result from instream turbidity; Figure 2.3 is an example of the
latter from a Montana river. In 2007, a tributary upstream of the Figure 2.3 data-collection site had a summertime high flow event and discharged highly turbid, sediment-laden water to the river. Productivity, measured indirectly as the magnitude of diel DO oscillation in the river, was notably dampened during the event (days 1 through 3) after which productivity increased as turbidity fell to more normal levels. Note the very high turbidity levels (> 600 NTU) that were needed to induce the dampening effect (Figure 2.3). Turbidity levels in western Montana streams are, outside of runoff, typically well below 50 NTU, and are not likely to limit instream productivity. In contrast, turbidity levels in eastern Montana prairie streams are sometimes above 100 NTU and midsummer values as high as 1,830 NTU have been measured in ones considered to have few human impacts. Turbidity in these prairie streams correlate well with suspended sediment, therefore it is likely that Montana prairie streams experience periodic reduction of their productivity due to sediment turbidity.

Stream velocity and high-flow events interact with eutrophication and influence benthic algae in two distinct ways. Water velocity — up to a point — can allow larger algae mats to grow than is possible in quiescent water because the flow induces nutrients to reach algae cells at the base of the mat which might otherwise be starved of nutrients by the algae growing above them. In contrast, high flow events beyond that which the algae are adapted to leads to reduced biomass via sloughing and scouring. Benthic algae scouring can occur in Montana during the summer due to isolated high flow events caused by localized thunderstorms; this is especially true in eastern Montana where such storms are common and stream flows tend to be flashy. High flow events act as a reset mechanism, after which rapid re-growth of the algae can and often does occur.
Grazing by fish and macroinvertebrates affects plant productivity in streams, however studies in this area of stream ecology are often conflicting and there is still much scientific debate. An excellent literature review is provided by Steinman 46, but only the most pertinent aspect of his review (biomass-grazing relationship) is summarized here. Most studies show — not surprisingly — that algal biomass decreases in response to fish and macroinvertebrate grazing, although a few studies show algae biomass actually increases. Algal biomass might not decline due to grazing because (1) grazer density and consumption rates are insufficient to induce a decline, (2) the type of grazing is not well matched to the dominant algae forms, and (3) other resources (e.g., nutrients) are limiting and biomass is low regardless of grazing. We are not aware of any published studies of grazing effects on aquatic plants carried out in Montana, and can only offer our own observations on this topic. Based on many years of working in Montana streams, it is our conjecture that grazing by fish and macroinvertebrates does not play a large role in reducing heavy algal growth in Montana streams. Snails are a common grazer 46, but we have not often observed high snail densities in either western or eastern Montana streams, probably due to harsh conditions in winter. An exception to this is in a few streams where the invasive New Zealand mud snail (Potamopyrgus antipodarum) has become established. Spring creeks also tend to have somewhat elevated snail densities. In near-pristine streams, biomass is probably constrained by low nutrients (reason No. 3 above). In other, more eutrophied streams, where we have observed very high algal biomass (often Cladophora) growing in thick mats, low
grazer density, or the less-than-preferable nature of the algae as a food source, seem equally likely to explain why grazing has not observably diminished heavy benthic algae growth. Alternatively, benthic algal biomass might increase due to macroinvertebrate grazers, because grazers can stimulate *Cladophora* growth by removing epiphytes growing on the surface of the algal filaments\(^{47,48}\).

Finally, there is inter-specific competition among stream aquatic plants. Of particular relevance in Montana is the competition observed between benthic plants and phytoplankton in the state’s eastern prairie streams. Benthic algae and phytoplankton compete with each other for resources (light, nutrients, etc.) and, once one or the other of the two plant groups gains the upper hand, a positive feedback loop can ensue that leads to the near domination by that group\(^{49}\). This process appears to occur in Montana prairie streams, most often when they become intermittent\(^{41}\). An analogous regional example of this process described in the scientific literature is found in shallow lakes of the Canadian Boreal Plain\(^{50}\). These lakes switch status from year to year between (state 1) phytoplankton domination and turbid water, and (state 2) domination by submerged aquatic vegetation and clear water. Harsh winters (i.e., non-biological factors) tend to reset the lakes each year, and so the state (1 or 2) that dominates the following year is dependent on small perturbations present in the spring\(^{50}\). As mentioned, a similar phenomenon has been observed in Montana prairie streams\(^{41}\)(see also Discussion, Appendix A), and where prairie streams are eutrophied this can result in heavy blooms of phytoplankton with chlorophyll \(a\) concentrations as high as 515 \(\mu g\) Chl \(a/L\)\(^{41}\). Phytoplankton influence DO as do benthic plants, and in eutrophied prairie streams low DO seems equally likely to result from phytoplankton blooms or from heavy benthic plant growth.
Section 3.0
Existing vs. Proposed Approach to Controlling Stream Eutrophication

3.1 The Push for Numeric Nutrient Criteria at the National Level

Eutrophication has long been recognized as a major water quality problem by EPA, illustrated by the fact that the agency undertook a national eutrophication survey of streams just shortly after its creation in the early 1970s. In the late 1990s EPA announced that all states and tribes must develop nutrient criteria for their respective waters, and by 2000 EPA had published a series of regionally-based numeric nutrient criteria recommendations. As of this writing, EPA’s current policy position is that each state and tribe should adopt numeric criteria, but they can develop and adopt the criteria according to mutually agreed upon plans and schedules. There are already a number of states that have adopted numeric nutrient criteria (e.g., Tennessee, Hawaii, Connecticut), and many more states are well along in the development process.

3.2 How Eutrophication has been Addressed in Montana Up Till Now, and How Numeric Nutrient Criteria Will Improve the State’s Existing Water Quality Standards

Although DEQ is developing statewide numeric nutrient criteria for the first time, eutrophication has long been recognized as a problem and water quality laws exist to help address it. How have the negative aspects of stream nutrient enrichment been addressed in Montana to date? Montana has several water quality standards that generally apply to eutrophication, including a limited number of numeric nutrient criteria. Numeric nutrient criteria are established on large reaches of the Clark Fork River and define reach-specific nutrient concentrations and benthic algae biomass levels for the river (ARM 17.60.631). These were adopted in 2002 and are intended to prevent nuisance growth of benthic algae by limiting the river’s N and P concentrations. Other numeric standards that are applicable to all state waters and that address eutrophication-related water quality problems are the numeric DO criteria, and the pH criteria (e.g., ARM 17.30.623[2][c] ). These standards are intended to protect fish and aquatic life uses.

In addition to the regulations cited above, Montana has a narrative criterion that covers other unwanted aspects of eutrophication. Narrative criteria are codified statements that describe, in a concise way, a water quality condition that must be maintained. However, unlike numeric criteria, there are no quantitative values associated with narratives. The Montana narrative criterion that is applicable to eutrophication specifies that “State surface waters must be free from substances attributable to municipal, industrial, agricultural practices or other discharges that will create conditions which produce undesirable aquatic life” (ARM 17.30.637[1][e]). Narrative criteria have the advantage that they are flexible and cover many potential situations (even unforeseen ones), but because they lack specificity, they are open to varied interpretations.

An obvious question that arises is “why adopt statewide numeric nutrient criteria if Montana already has other criteria that address eutrophication?” DEQ asked this very question when it began to develop statewide numeric nutrient in 2000 and concluded that the existing criteria were
3.0 Existing vs. Proposed Approach to Controlling Stream Eutrophication

not sufficiently addressing some types of water quality problems\textsuperscript{53} (Clark Fork River criteria excluded). Clearly, something about the DO, pH, and narrative criteria was not and is not working when it comes to stream eutrophication, since eutrophication problems continue to be common in Montana\textsuperscript{54}. When it comes to eutrophication, the shortfall of the numeric DO and pH criteria is that they are \textit{effect}, or secondary, variables in streams, and require one to further seek the specific, primary cause of the impact. That is, DO and pH are being driven by other factors and those “other factors” are often excess nutrients. To properly implement the DO criteria where eutrophication is involved requires nighttime measurement of DO (when the minima usually occur), a cause-effect linkage between DO and nutrients, and an understanding of the nutrient concentrations that would prevent the low DO from occurring. Similarly, the pH criteria require an understanding of a waterbody’s natural background pH, the degree of change from background, and the cause. Thus, if one knew the nutrient concentrations that could prevent exceedences of the DO and pH criteria in a waterbody, one has a good chance of actually attaining the DO and pH criteria because the root cause of the problem would be addressed. That is exactly what numeric nutrient criteria are intended to do.

The narrative criterion (ARM 17.30.637[1][c]) has more difficult implementation challenges than the DO and pH criteria do. In particular, there are no definitions in rule of what “undesirable” aquatic life is, or, if that could be determined, what the levels of this aquatic life should be held to. If undesirable aquatic life can be defined and maximum allowable levels of it are established, then the situation resembles that of DO and pH in that one needs then to determine appropriate nutrient concentrations where enrichment is involved. But because undesirable aquatic life has not been defined heretofore, the application of this criterion has been subject to individual interpretation and, consequently, debate. In developing the numeric nutrient criteria it was necessary to (1) identify and quantify undesirable (i.e., nuisance) aquatic life attributable to eutrophication, (2) determine a level of that aquatic life that would not harm the beneficial uses, and (3) identify the nutrient concentrations that would maintain said aquatic life at non-nuisance levels. As a result, the numeric nutrient criteria in this document closely reflect the spirit and intent of the narrative criterion but also provide sufficient detail to make it of practical value.

3.3 Recommended Nutrients, and Determining Compliance with the Criteria

Criteria for total N (TN), total P (TP), and nitrate + nitrite (NO\textsubscript{2+3}) are proposed. Research shows that total nutrients provide better overall correlation to eutrophication problems in streams than do soluble nutrients\textsuperscript{55-57} and, in addition, EPA has indicated that TN and TP are the minimum acceptable nutrient criteria\textsuperscript{58}. Although total nutrients correlate best overall with eutrophication, soluble NO\textsubscript{2+3} concentrations are also important. This is because there is the potential that if TN criteria are adopted in the absence of accompanying NO\textsubscript{2+3} criteria, the water quality benefits desired would not be achieved. For example, if a point source discharge met the TN criterion but the discharge was almost all nitrate, water quality problems would in all likelihood continue. This is particularly true in eastern Montana prairie streams, where there are strong indications of nitrogen limitation and studies suggest that special attention should be paid to NO\textsubscript{2+3}\textsuperscript{59,60}. We believe that soluble P criteria need not be promulgated because, as will be shown in Section 6.3.2, the TP criteria concentrations being recommended should maintain...
3.0 Existing vs. Proposed Approach to Controlling Stream Eutrophication

soluble P at acceptable levels. Therefore, we recommend TN, TP, and NO$_{2-3}$ for numeric nutrient criteria.

Recommendations are provided for determining compliance with the numeric nutrient criteria in Appendix H (for TMDLs and 303(d) listing) and Appendix I (for point sources) in a recently completed technical report$^{61}$. Readers interested in the details of the recommended compliance-determination techniques should read the aforementioned appendices. The most important point to summarize here is that the criteria will have a certain allowable exceedance rate based on appropriate statistical evaluation of the data. Specifically, 20% of the data from a population of water quality data could exceed any given criterion and still be in compliance with the standard. This exceedence rate was determined empirically using Montana data and falls within the range of exceedence rates recommended by EPA$^{61}$. A 12 sample minimum is recommended.

3.4 Difficulties with Numeric Nutrient Criteria

Numeric nutrient standards will help Montana to better protect beneficial uses and water quality. Numeric nutrient criteria are not without their own difficulties, however. In our opinion, the major issues are (1) the question of geographic specificity of the criteria, and (2) the achievability of the criteria using current wastewater treatment technologies.

Nutrient criteria concentrations proposed by DEQ will be different from place to place due to local differences in geospatial features (e.g., climate, geology, soils) and their combined effects on stream nutrient concentrations. DEQ has carried out substantial research to assure that the classification system used to differentiate nutrient expectations is regionally appropriate (much more on this, Section 4.0). But whenever a classification system is used, decisions have to be made as to whether it is better to lump individual categories, or to further split them. There are two ends to this continuum; at one end, there is a classification system so “split” that each stream falls in its own class (i.e., each stream is unique), while at the other end of the spectrum all streams are “lumped” together (i.e., all streams are the same). Neither of these is appropriate for nutrient criteria. Instead, DEQ strove to find a useful and practical balance between these two extremes when developing the numeric nutrient criteria.

Nevertheless, there will be cases where the criteria are not exactly appropriate for a given stream due to local conditions not sufficiently addressed by the classification system or because the classification is too coarse. Some of the localized, confounding environmental factors that change the way eutrophication is manifested in streams have already been presented (high flow events, shading, etc.; Section 2.3). The confounding environmental factor we consider most likely to render the criteria inappropriate is the near-field (i.e., proximate) effect of upstream dams (which are addressed by other MT laws, e.g. 75-5-306(2), §MCA), and upstream natural lakes. However, there are other state water quality laws (e.g., “it is not necessary that wastes be treated to a purer condition than the natural condition of the receiving stream...”, 75-5-306(1), §MCA) not discussed in this document that can address any gross criteria misfits on a case-by-case basis. It is worth pointing out that the establishment of beneficial uses and water quality criteria in Montana have traditionally been somewhat “broad-brush”. This approach has the advantage that all waterbodies are protected under law, but as a consequence criteria or use
classes may not make sense for some specific waterbodies and warrants additional site-specific consideration.

The other major issue that has become clear is that nutrient concentrations that prevent the unwanted aspects of eutrophication are quite low relative to current wastewater treatment technologies. Scientific studies show that it only takes small amounts of nutrient enrichment to manifest changes in streams\(^{43,62}\); this region of the country appears to be particularly sensitive and the specifics of this will be detailed in Section 6.0. The implication for Montana is that the criteria will be difficult to achieve in some places, especially where a point-source discharge is a large proportion of a receiving stream’s volume. DEQ is developing implementation policies that will help dischargers deal with this; those efforts are on going, and we will not present those here. It is also important to note that wastewater technologies are rapidly advancing, hence, lower and lower N and P concentrations can be routinely achieved for less money. DEQ anticipates that the numeric nutrient criteria are, ultimately, achievable, even if dischargers need time for treatment technologies to mature and costs to come down.
Section 4.0
How DEQ Selected a Geographic Stratification System to Apply Different Nutrient Criteria in Different Places

An essential step in setting numeric nutrient criteria involves deciding how to divide the state into regions or nutrient zones in which a single N or P criterion would apply. This section discusses why such a geographical stratification system is necessary and how it was developed.

4.1 Purpose of Developing a Geographic Stratification System

As mentioned in Section 3.4, one approach to setting numeric nutrient criteria would be to identify a single nutrient concentration (e.g., total P) that is protective of beneficial uses and apply that concentration as a uniform criterion across the state. Such a single state-wide numeric criterion approach would be, however, deficient for several reasons.

The natural sources of N and P in surface water are mainly geology, soils, and vegetation\textsuperscript{35, 63}. Climactic conditions, and other regional variables, may affect the rate at which N and P are released to the surface waters from these sources. But different regions of the state are endowed with different types of soil, different kinds of underlying geology and experience differing climates. As a consequence, nutrient concentrations in rivers and streams are expected to show a natural variability across the state even in the absence of impairment from human activities. It would be entirely natural for some regions to manifest higher background levels of nutrient concentrations than other regions. A single, state-wide criterion therefore has the serious disadvantage that it may either (1) require the attainment of nutrient levels that are below the natural background level for a region (imposing unnecessary and unrealistic attainment costs on local communities) or (2) allow the build-up of nutrient concentrations that are above an acceptable background level for that region (possibly leading to the problems associated with eutrophication described in Section 2.2).

Furthermore, the association between nutrient concentrations and beneficial uses involves regional factors. Owing to the regional variability of plant and animal species and their differing ability to adapt to environmental conditions, a nutrient concentration that does not compromise beneficial uses in one region may indeed affect beneficial uses significantly in another. Therefore, the very process of determining a single concentration that does not compromise beneficial uses may be misguided and inaccurate if applied blindly from one region to another.

For these reasons, a more scientifically accurate approach to setting nutrient criteria would require partitioning the state into zones that share common characteristics such as comparable background nutrient concentrations in their surface waters. One way to do this would be to divide the state into areas which share the same basic soil, geology, vegetation, climate, and regional topographical features. In each of these zones, a single criterion for each nutrient can be reasonably applied because it is reasonable to conclude that the background concentration within each zone is similar, and that the beneficial uses, expected relationships between causal (nutrient) and response variables (e.g., benthic algae growth), and effect thresholds would be comparable.
4.0 How DEQ Selected a Geographic Stratification System to Apply Different Nutrient Criteria in Different Places

4.2 Conceptual Approaches for Developing a Geographic Stratification System

There are potentially an infinite number of ways in which the state could be divided, or stratified, into nutrient zones. The most effective stratification methodology, however, would be the one that maximizes the difference in concentration between zones and minimizes the variance within zones. The specific statistical tests which can be used to measure the performance of a proposed stratification methodology are described in Section 4.5.

DEQ began by considering three conceptual approaches for determining the nutrient zones:

- A purely empirical approach based on iterative random delineation of geographical strata followed by statistical analysis of stream monitoring data;
- An empirical approach based on statistical methods known as factor analysis which uses stream monitoring data together with a host of regional environmental variables to create geographic strata; and
- A combined approach based on extant (i.e., existing) geographic classification systems which are subsequently verified and refined by the analysis of empirical data from within the proposed extant zones.

These approaches require the availability of a database of observed nutrient concentrations evenly sampled from all areas of the state from streams classified as “reference” or “background” streams. As will be discussed in greater detail in Section 6.2, reference sites represent our best approximation of stream condition in the absence of noteworthy human disturbance or alternation, although they are not all pristine.

The data requirements of the first two approaches are likely to be much greater than the third approach. The potential approaches to delineating nutrient zones are discussed in further detail in Sections 4.2.1, 4.2.2 and 4.2.3 below. The preferred approach is identified in Section 4.2.4.

4.2.1 The Random Delineation Approach to Establishing Regional Nutrient Zones

This approach is based on repeatedly dividing the state into random nutrient zones and then testing to determine which of the iterations maximize inter-zone nutrient concentration differences and minimize intra-zone differences. To use an analogy, the random delineation of nutrient zones may be thought of as equivalent to splashing several different colors of paint onto a map of the state, assuring that none of the color blobs overlap. Each color would represent a nutrient zone. This process is repeated a large number of times. The map resulting at the end of each iteration would represent one possible classification scheme for nutrient zones. Each resulting map would have to be evaluated using the available database to assess its performance as a stratification methodology. The map in which nutrient zones are most effective at

November, 2008

16
maximizing inter-zone differences and minimizing intra-zone differences would be selected as the best scheme for delineating nutrient zones.

The most significant advantage of this empirical approach is that it has the potential to generate a nutrient zone classification system which performs better at maximizing inter-zone variance and minimizing intra-zone variance than a system generated by any alternative method for a given dataset. In other words, through the sheer power of number-crunching this technique can come up with the best classification scheme for nutrient zones as measured by statistical performance metrics.

The empirical approach based on random delineation has several drawbacks. Most importantly, it does not take advantage of any established theories of water quality. The exclusive focus on numerical analysis means that the results will only be as good as the available database — this could result in high levels of uncertainty and misleading results. Another disadvantage is that the empirical approach is computationally intensive. Also, the size and shapes of the resulting zones could be impractical to implement.

4.2.2 The Factor Analysis Approach to Establishing Regional Nutrient Zones

The factor analysis approach attempts to identify the regional variables which best explain the variance in background nutrient concentrations using complex statistical methods broadly known as factor analysis. Factor analysis is a term encompassing a family of statistical techniques which reduce multiple variables (such as geology, climate, etc) into smaller groups to minimize the variance of the dependent variable (in this case nutrient concentrations) within these groups. In the context of nutrient concentrations, the groups resulting from a factor analysis can effectively be thought of as a basis for defining nutrient zones. Unlike an extant classification system (described next), factor analysis combines regional variables without reference to theoretical concepts from water quality science.

The advantage of this approach is that it is less computationally intensive than the empirical random delineation method described above. A second advantage is that it seeks to determine associations and correlations between the dependent and independent variables without reference to any existing classification system; it is possible that this approach could therefore reveal associations overlooked by standard nutrient theory.

The disadvantage of the factor analysis approach is that is will only be as good as the database upon which it operates. Errors and imbalances in the database will result in inaccurate nutrient zones and high levels of uncertainty. The methods of factor analysis are also statistically complicated and would require a high input of analytic effort. Nutrient zones based on factor analysis may be difficult to communicate to the public and complicated to implement.

4.2.3 The Combined Approach to Establishing Regional Nutrient Zones

The combined approach to establishing nutrient zones endeavors to take advantage of theoretical knowledge of water quality science as well as empirical data on nutrient concentrations. As described above, theoretical considerations lead to the expectation that areas with similar soils,
4.0 How DEQ Selected a Geographic Stratification System to Apply Different Nutrient Criteria in Different Places

geology, climate and other regional environmental factors are likely to have similar background nutrient concentrations in their surface waters. In the combination approach to establishing nutrient zones, several potential extant classification systems are proposed. The extant classifications are selected based on their likelihood of properly reflecting the natural variability of stream nutrients concentrations. These alternative extant classification systems are then evaluated based on the available empirical evidence to determine which performs best in terms of maximizing inter-zone variance and minimizing intra-zone variance. Thus, the combination approach is a dual approach comprising two distinct steps:

Step 1: Propose alternative extant classification schemes that, based on theoretical considerations from water quality science, are likely to be good candidates for segregating stream nutrient concentrations.

Step 2: Evaluate the performance of the proposed classification schemes using empirical data to determine which scheme is best able to maximize inter-zone variance and minimize intra-zone variance; statistical tests will also ascertain whether the proposed scheme is statistically significant.

The primary advantage of the combined approach is that it seeks to balance theoretical knowledge from water quality science with empirical field observations in the delineation of nutrient zones. The proposed extant classification schemes are expected to stratify the state into zones with similar background nutrient concentrations in surface waters based on well established theoretical principles of science. The empirical analysis which follows will then confirm whether the proposed schemes do achieve this objective and will identify the best performing stratification methodology amongst the options analyzed. Thus, the approach is not overly subject to uncertainty and error resulting from errors or imbalances in the available water quality database.

A potential disadvantage of the combination approach is that the final result for a stratification methodology will only be as good as the proposed alternatives. That is, the approach will choose the best of a given set of alternatives. It is possible that the best possible approach may not have been included amongst the set of extant classification systems tested.

4.2.4 DEQ’s Preferred Approach

Based on the advantages and disadvantages described above, DEQ decided against pursuing either the random delineation or the factor analysis for determining nutrient zones because of the high data requirements, statistical uncertainties, computational intensity, and analytic complexity. DEQ instead favored the combined approach (Section 4.2.3), which builds on theoretical concepts of water quality science using rigorous empirical analysis.

4.3 Extant Classification Systems

The first step in using the combined approach is to identify potential extant classification systems for nutrient zones. When choosing amongst potential extant classification systems, one of DEQ’s choices was Omernik ecoregions. Designed to serve as a spatial framework for
environmental resource management, ecoregions denote areas within which ecosystems (and the
type, quality, and quantity of environmental resources) are generally similar. The ecoregion
concept is based on the premise that ecologically similar regions can be identified through
analysis of the patterns and composition of biotic and abiotic factors that affect or reflect
differences in ecosystem quality and integrity. These factors include geology, physiography,
vegetation, climate, soils, land use, wildlife, and hydrology. James Omernik carried out a
national eutrophication survey of streams in the 1970s and, following up on that work,
developed the first of his ecoregion maps. The stated purpose of the ecoregion maps was to
classify streams for more effective water quality management. Thus, from the outset,
ecoregions were developed and designed for making decisions about streams and their water
quality.

A Roman numeral classification scheme has been adopted for different ecoregion scales. Level I
is the coarsest level, dividing North America into 15 ecological regions, whereas at level II the
continent is subdivided into 52 classes (Available on the web at:
http://www.epa.gov/wed/pages/ecoregions.htm). Level III and level IV are the hierarchical
levels evaluated as part of this analysis. Montana contains parts of 7 level III ecoregions and
also 85 level IV ecoregions (Figure 4.1).

In addition to examining the usefulness of ecoregions as a classification system for establishing
nutrient zones, DEQ also evaluated two other extant classification systems; Strahler stream
order, and lithologic groupings (i.e., underlying geology). Strahler stream order groups
streams of a similar dimension and flow, and had the potential to delineate meaningful nutrient
zones if in fact nutrients in streams were largely a function of upstream watershed area.
Lithology groups land areas with a similar underlying geology, and since the rocks through
which streams flow are known to influence their water chemistry (including nutrients, e.g.
nitrogen), lithology was considered to be a good candidate for nutrient zones.
4.0 How DEQ Selected a Geographic Stratification System to Apply Different Nutrient Criteria in Different Places

4.4 Historical Database of Montana Stream Monitoring Data

The extant classification systems proposed in Section 4.3 needed to be verified and supported by statistical tests using empirical data. The main data source for the analyses was from the U.S. EPA’s Storage and Retrieval (STORET) database. The database was supplemented with all river and stream nutrient data from DEQ found in modernized STORET, which were collected from 2000 to 2004. Additional data sources included Montana river and stream data collected by the University of Montana, Utah State University, the U.S. EPA’s Environmental Monitoring and Assessment Program (EMAP), and reference-stream nutrient data up through 2007. Greater detail on the database and the quality control measures used to assemble it are on page 454-456 of Supplee et al.\(^6\). The database contained approximately 13,000 sampling sites and over 140,000 total records. In the following summaries of data analysis, the term “general population” refers to all observations from both reference and non-reference sites. The term “reference population” refers to nutrient data only from reference sites (reference sites are further discussed in Section 6.2).
4.0 How DEQ Selected a Geographic Stratification System to Apply Different Nutrient Criteria in Different Places

4.5 Statistical Methods used to Test the Geographic Nutrient Zones

The extant classification systems proposed in Section 4.3 can be used to generate different combinations of stratifying and sub-stratifying hierarchy for the creation of nutrient zones. For example, one hierarchy could define zones based on each Strahler stream order within level III ecoregions. An alternative potential hierarchy could define zones based on each level IV ecoregion within each level III ecoregion.

With several potential stratification hierarchies, statistical tests are necessary to identify the best-performing hierarchy and to confirm that the proposed hierarchy is indeed statistically significant. As described earlier, the best performing hierarchy is one that stratifies the state into nutrient zones such that the inter-zone variability of nutrient concentrations is maximized and the intra-zone variability of nutrient concentrations is minimized; the hierarchy must also be statistically significant and not just the result of sampling variance.

Non-parametric and parametric statistical methods were both used to examine whether the proposed extant classification hierarchy did indeed result in nutrient zones in which nutrient concentrations were different from one another at an adequate level of statistical significance. A test was considered statistically significant if we could deduce with 95% confidence that the observed differences were indeed reflective of true differences in the data groupings and not the result of sampling variance. Parametric tests were preferred if the distributional requirements of the underlying data were satisfied; the results of parametric tests provided more information than non-parametric tests especially for examining the statistical significance of sub-stratifying methodologies. The non-parametric tests have the advantage, however, of not requiring the underlying data to adhere to any particular statistical distribution. If more than one proposed classification system is found to be statistically significant, it is possible to investigate which classification maximizes inter-zone variability by assessing which classification results in the most nutrient zones in which nutrient concentrations are different from one another at an adequate level of statistical significance. These tests are discussed in Section 4.5.1 and 4.5.2.

Statistical measures were also designed to measure the performance of the alternative extant classification methodologies at reducing intra-zone variance. These tests are discussed in Section 4.5.3.

4.5.1 Non-Parametric Tests for Determining Differences Between Stratified Populations

A stratification methodology may be considered statistically significant if there are differences in nutrient concentrations between the zones defined by the methodology, i.e., if at least one zone may be considered to have a higher or lower median concentration than the other zones. In order to test for statistically significant differences between the median nutrient concentrations of different strata within a given stratification hierarchy, we used the non-parametric Kruskal Wallis test. This test is very similar to a one-way ANOVA in which the data are replaced by their ranks. The main advantage of the Kruskal Wallis test is that it does not require the populations to be normally distributed although it does assume that the data in each grouping follow a similarly shaped distribution. The Kruskal Wallis test is an extension of the Mann-Whitney test (also known as the Wilcoxon Rank sum test) to three or more data groupings. The test was used only
on the median database. (As described above, the median database is a version of the water quality database in which each sampling station is represented by a single median value for each season for each nutrient grouping. The median database reduces imbalances in sampling frequency and is likely to obey the distributional assumptions of the parametric tests.)

A 95% confidence level was used to identify statistically significant differences. If the test indicated the existence of statistically significant differences in median concentrations between the strata, a post-hoc non-parametric multiple comparison test was implemented\textsuperscript{68}. These procedures helped determine which strata could be considered different from one another.

### 4.5.2 Parametric Tests for Determining Differences Between Stratified Populations

We used analysis of variance (ANOVA) to test for statistically significant differences between the mean nutrient concentrations of different strata for a given stratification methodology. Although similar to the non-parametric Kruskal Wallis test, ANOVA offers the substantial advantage of being able to test for the statistical significance of sub-stratifying methodologies. ANOVA procedures are most accurate when the underlying populations are normally distributed with equal variance in each stratum. For our analysis, tests showed that ANOVA results could be considered robust to the observed levels of non-normality and inequality of variance (see page 11, Varghese and Cleland\textsuperscript{69}).

ANOVA was implemented only on the median database because this database reduced imbalances in sampling frequency and better obeyed the distributional assumptions of the parametric tests. A 95% confidence level was used to identify statistically significant differences. If the test indicated statistically significant differences in mean concentrations between the strata, a post-hoc parametric multiple comparisons of means was performed using the Bonferroni adjustment. In order to test the statistical validity of sub-stratification, we used a nested ANOVA model with sub-strata nested within the main strata. We then used the Wald test to test the significance of the sub-stratification term in the nested model. The Wald test is a way of testing the significance of particular explanatory variables in a statistical model, including nested variables.

The coefficient of determination, represented as $R^2$, is the proportion of the total variability in the dependent variable that is accounted for by the model. It is an indicator of the goodness of fit of a model. An $R^2 = 1$ indicates that the model accounts for all the variability of the values of the dependent variables in the sample data. At the other extreme, an $R^2 = 0$ implies that the model explains none of the variability. This measure must be used with caution. Statisticians warn that a high $R^2$ does not assure a valid relation just as a low $R^2$ does not mean the model is without value.

The $R^2$ and adjusted $R^2$ statistics were computed for all ANOVA runs in our analysis. While indicative of a model’s goodness of fit, these measures should not be used alone to select between alternative statistically valid stratification methodologies, because adding variables or sub-strata to a model will always improve the $R^2$ measure. Instead, once a set of statistically significant stratification methods have been determined, the selection of the optimal method
should also consider a priori ecological, biological, and hydrogeologic considerations, and practical ease of applicability.

### 4.5.3 Computation of Measures of Variance for Alternative Stratification Methods

Section 4.5.1 and 4.5.2 discussed tests to ascertain whether statistically significant differences existed between proposed nutrient zones. To assess the performance of alternative stratification methodologies in minimizing variation within nutrient zones, two measures of intra-zone variance were computed: the mean coefficient of variation and the coefficient of efficiency.

For each stratification methodology, the mean coefficient of variation (MCV) was computed as follows, based on a definition provided in Robertson et al.⁷⁰:

\[
MCV = \sqrt{\frac{\sum (CV^2 \times n)}{N}}
\]

\[
CV = \frac{\text{StDev}}{\bar{X}}
\]

where,

- \( CV \) is the coefficient of variation of each group (or area);
- \( n \) is the number of observations in each group;
- \( N \) is the total number of observations in all of the groups;
- \( \text{StDev} \) is the standard deviation of each group; and
- \( \bar{X} \) is the mean concentration of each group.

A shortcoming of the MCV measure is that it is likely to improve (i.e., show lower absolute values) with increasing stratification. Therefore it would only be appropriate to use the MCV to assess the performance of alternative stratification schemes if the schemes divide the state into roughly equal numbers of strata.

Hydrologists have proposed the coefficient of efficiency as a means of evaluating the goodness-of-fit of hydrologic and hydroclimatic models⁷¹. This measure is defined as follows:

\[
COE = 1 - \frac{\sum_i (O_i - P_i)^2}{\sum_i (O_i - \bar{O})^2}
\]

where,

- \( O_i = \) Value of the \( i^{th} \) observation
- \( P_i = \) Predicted value corresponding to the \( i^{th} \) observation (equal to the mean of the observations in the stratum of the \( i^{th} \) observation)
- \( \bar{O} = \) Grand mean of observed values
Thus, the COE in this analysis will equal the ANOVA $R^2$. This measure can vary from minus infinity (poor model) to 1.0 (perfect model). Like the MCV, the COE has the shortcoming of being likely to improve (increase) with increasing stratification. Therefore it would only be appropriate to use the COE to assess the performance of alternative stratification schemes if the schemes divide the state into roughly equal numbers of strata.

Although the MCV and COE will usually be negatively correlated (i.e., high MCV associated with low COE and vice versa), there may be exceptions to this trend. These exceptions may occur because the MCV is weighted by the number of observations in each group and because the COE is more sensitive to departures from the grand mean.

4.6 Other Issues Affecting the Choice of Nutrient Zones

Although the chosen approach and the supporting statistical analysis may suggest the adoption of a particular system of geographical stratification, there are other pertinent issues affecting the choice of nutrient zones. These issues include the optimal number of nutrient zones, whether specific zones are required for specific nutrients, and whether there is the need for season-specific nutrient zones. These issues are discussed in greater detail in this section.

4.6.1 The Ideal Number of Regional Nutrient Zones

In developing any geographic stratification methodology an important issue that must be addressed is the appropriate scale and therefore number of nutrient zones. The disadvantages of a single zone, or too few zones, have been described. Too many zones are undesirable because they require extensive data; without a sufficient number of empirical observations from each proposed zone, the numeric nutrient criterion for that zone will suffer from a high degree of statistical uncertainty. And large numbers of zones will likely lead to excessive regulatory complexity. Deciding on an appropriate number of zones is a matter of regulatory judgment which must balance the need for accuracy in region-specific nutrient criteria vs. issues of sample size and regulatory complexity. One mechanism to reduce the number of nutrient zones in an extant sub-stratification methodology is to single out only those zones that are empirically determined to have different average nutrient concentrations from their parent stratification. For example, a methodology based on substratifying level IV ecoregions within level III ecoregions would require the formation of 34 nutrient zones in the Middle Rockies ecoregion since there are 34 level IVs nested within the Middle Rockies ecoregion in Montana. Multiple comparisons (based on procedures to compare nutrient means or medians, such as the t-test, for instance), however, could be used to compare the concentrations in each level IV ecoregion to the combined concentration in the remaining 33 ecoregions of the Middle Rockies to see if they are significantly different. Such comparisons may reveal that only particular level IV ecoregions have sufficiently distinct concentrations from the average level in the Middle Rockies to warrant separating them out. The analysis just described were carried out, and a number of level IV ecoregions were found to be unique relative to their parent ecoregion61. These results will be incorporated with other considerations relating to level IV ecoregions, and discussed again in Section 7.4.
4.0 How DEQ Selected a Geographic Stratification System to Apply Different Nutrient Criteria in Different Places

4.6.2 Nutrient-Specific Zones

Nutrient zones applicable for a single nutrient group (e.g., TP) may not match spatially with nutrient zones derived for another nutrient group (e.g., TN). Regulators then have to decide whether they wish to delineate separate nutrient zones for different nutrient groups, as has been proposed elsewhere\(^{70}\), or use nutrient zones that apply to multiple nutrients. The advantage of nutrient specific zones is that they are more likely to produce the most desirable ecological outcomes. The disadvantage of nutrient specific zones is increased regulatory complexity and potential public resistance.

4.6.3 Seasonal Considerations

Ecoregions are likely to show distinct inter-seasonal nutrient concentrations and, therefore, stratifying the nutrient data by different seasons would further improve the characterization of regional nutrient concentrations. Flowing waters often demonstrate distinct seasonal nutrient concentration patterns\(^{72}\). For example, P is frequently associated with total suspended sediment\(^{73}\) and during spring runoff in streams both TSS and TP can be orders-of-magnitude higher than at other times\(^{74}\). Seasonal variation in stream nutrient concentrations is not only influenced by abiotic factors such as runoff patterns, but also by biological uptake and release by organisms such as aquatic plants. Aquatic plant growth — including algal growth — is influenced by (among other things) light availability and temperature, which are themselves climatically driven. Therefore, stratifying the nutrient data seasonally to better characterize nutrient concentrations across time had to consider not only hydrologic patterns, but also climatic factors such as the onset of cold winter temperatures.

Giving consideration to the factors above, an analysis was completed and three seasons were defined for each ecoregion: a growing season, which would roughly correspond to the summer months; winter, which would follow the growing season; and runoff, which would terminate the winter period and comprise the yearly high flow period\(^{60}\). Some minor changes to the start and end dates were carried out since the publication of the aforementioned work, and these are shown in Table 4.1. Table 4.1 provides the current recommendations for start and end dates for the winter, runoff and growing season for each level III ecoregion.
4.0 How DEQ Selected a Geographic Stratification System to Apply Different Nutrient Criteria in Different Places

Table 4.1  Start and Ending Dates for Three Seasons (Winter, Runoff and Growing), by Level III Ecoregion.

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

4.7 Results of the Empirical Analysis

As part of the process of creating regional nutrient zones, the proposed extant stratification systems were subjected to various statistical tests (described in Section 4.5) using the historical nutrient database of stream monitoring data. The results of these tests are summarized below:

- Of the various coarse scale stratification systems tested (level III ecoregions, Strahler stream order, and geology), level III ecoregions produced strata that differed significantly from one another in terms of their median nutrient concentrations, for the reference data, for all nutrient groups (TKN, TN, NO₂⁺³, SRP, TP) except total dissolved P (TDP), which has a very small dataset. The efficacy of stratification by level III ecoregions is most apparent in the growing season, and for year-round data.

- Post-ANOVA Wald tests were used to verify the statistical significance of various sub-stratification methodologies of the coarse-scale strata, for the reference population, on a limited selection of nutrients. For year-round data, sub-stratification by level IV ecoregions was consistently an improvement over stratification by level III ecoregions. The other sub-stratification methods did not show statistically significant results. However, sample size was limited at this level of stratification for the reference population and the power of these tests is likely to be low. (Low statistical power results in an inability to declare a stratifying parameter as significant even though in reality it may be significant. Low power can result from inadequate sample size.)

- Analysis of the measures of intra-zone variance (MCV and COE) indicates that the statistically significant stratification methodologies are more successful in explaining variance for nitrogen-group nutrients than for phosphorus groups. The stratification methodologies have the most explanatory power in the winter season. The growing season appears to be the most noisy.

- The measures of intra-zone variance indicate a considerable improvement in the measures of variance with increasing sub-stratification. However, as explained earlier,
4.0 How DEQ Selected a Geographic Stratification System to Apply Different Nutrient Criteria in Different Places

this improvement may partly be the result of a fewer number of observations contributing to each stratum and should be considered with caution.

- When sub-stratifying level III ecoregions by level IV ecoregions, multiple comparison tests indicate that only certain level IV ecoregions are statistically different from their parent level III ecoregions. It is thus possible to reduce the overall number of nutrient zones by creating nutrient zones only for those level IV ecoregions that are in fact distinct from their parent level III ecoregion.

- Seasonal analysis of the stream data indicate that seasonal differences in background nutrient concentrations are significant when seasons are defined per the methodology outlined in Section 4.6.3.

If the reader would like further detail on these topics, further explanation is available in both documents by Varghese and Cleland61,69.

4.8 Conclusions about the Geographic Stratification Systems

Based on the results described above, the following conclusions were drawn about geographically stratified nutrient zones in Montana:

- Level III ecoregions and level IV ecoregions constitute statistically significant systems for stratifying the state for most nutrients in most seasons for both the general- and reference-data population.

- Nutrient zones based on sub-stratifying level III ecoregions by level IV ecoregions may be regarded as the best performing stratifying methodology examined in our analysis, based on tests of statistical significance, measures of variation, and a priori theoretical considerations (i.e., the underlying theoretical basis of the maps themselves).

- Most of the stratifying methodologies considered in this analysis perform better for nitrogen than for phosphorus. It may be possible to generate more complex and efficient stratifying methods specifically for the phosphorus-group nutrients. However, for simplicity, it is advisable to create common nutrient zones for N and P together, rather than different zones for each nutrient.

- When using level IV ecoregions as a sub-stratifying methodology within level III ecoregions, it is advisable to reduce the overall number of nutrient zones by identifying and creating separate zones only for those level IV ecoregions that are significantly different from their parent level III ecoregion. This will be discussed further in Section 7.4.
Section 5.0
Identifying Eutrophication Impacts on Sensitive Beneficial Uses

One of the most important aspects of setting criteria is determining when beneficial uses begin to become harmed. An example of harm to a beneficial use was given back in Section 1.4. In this section, we will address eutrophication-specific harmful effects on uses.

As noted in Section 2.2, heavy algae growth, especially by filamentous forms, is a very common effect of eutrophication and can be seen every summer in many Montana streams. It has been generally observed that as the level of benthic algae increases in a stream, the suitability of the waterbody for public recreation decreases\(^2,58,75,76,76\). To verify this observation, in 2006 DEQ and the University of Montana carried out a statistically rigorous statewide public perception survey concerning benthic river and stream algae\(^77\). (The study will shortly be published in the *Journal of the American Water Resources Association*.) A copy of the article can be requested from M. Suplee\(^c\). A web-available summary is at [http://www.umt.edu/watershedclinic/algaesurveypix.htm](http://www.umt.edu/watershedclinic/algaesurveypix.htm). Photographs of varying levels of stream-bottom algae, as seen in typical western Montana gravel-bottomed streams, were shown to Montana citizens and recreators on Montana rivers and streams across the state. Stream algae levels were quantified as chlorophyll $\alpha$ ($\text{Chl} \alpha$) and ash free dry weight (AFDW) per square meter of stream bottom (this information was not provided to survey participants; just the pictures). Participants were asked how the algae level shown in each photograph would affect their recreational use of the stream or river, whatever that recreation might be (e.g., swimming, fishing, boating, etc.). The results were remarkably clear. 70% or more of the public felt that algae levels less than or equal to 150 mg $\text{Chl} \alpha$/m$^2$ ($\leq 36$ g AFDW/m$^2$) were acceptable for recreation. But then a sharp threshold occurred, and only 30% or less of the public considered algae levels at or above the next level up — 200 mg $\text{Chl} \alpha$/m$^2$ (95 g AFDW/m$^2$) — to be acceptable for recreation. The more elevated algae levels in the survey were clearly viewed as undesirable aquatic life by the public majority. And, the sharp change in public majority opinion concerning the acceptability of benthic algae levels can be used to define the threshold where the recreation use becomes harmed. (As a frame of reference for the reader, the algae level in Figure 2.1 is 300 mg $\text{Chl} \alpha$/m$^2$, whereas the algae level in Figure 7.1 is $< 8$ mg $\text{Chl} \alpha$/m$^2$.)

Salmonid fishes are common in most western Montana streams. How does the algae level found to be a recreation impact threshold (150 mg $\text{Chl} \alpha$/m$^2$) relate to the ecology of these fish? Studies are few, but there is an excellent and applicable study from British Columbia in which N and P are added to a small, low-nutrient river during summer in order to observe the effects on benthic algae and salmonid fish\(^14\). The study shows that with increased N and P concentrations algae reach maximum average values of 150 mg $\text{Chl} \alpha$/m$^2$, filamentous algae become much more dominant, and juvenile salmonids show significant weight gain. Fish grew better probably because of increased macroinvertebrate abundance (i.e., more food). Thus, what is a harm threshold for one use (algae level of 150 mg $\text{Chl} \alpha$/m$^2$; recreation) equates to an enhancement of another use (salmonid fishes). Because fish benefit from some enrichment, does it make sense to allow eutrophication to proceed further, growing larger fish and, consequently, growing more algae? Not really. The initial benefits from nutrient enrichment are subsequently lost when too

\(^c\) E-mail: msuplee@mt.gov  Phone: (406) 444-0831

November, 2008  28
5.0 Identifying Eutrophication Impacts on Sensitive Beneficial Uses

much enrichment occurs (note in Figure 2.2, for example, how the number of salmonid fish per unit stream length declines again where TP > 0.06 mg/L)\textsuperscript{28,78}. Furthermore, it does not make sense to allow streams and rivers to be strongly eutrophied in order to grow more and larger fish if that action clearly results in an impact to stream recreation which, of course, Montana’s world-renowned fishing is a part of. That is, it might be possible to grow bigger trout at algae levels greater than 150 mg Chl\textsubscript{a}/m\textsuperscript{2}, but only a minority of the public (30% or less) would be interested in recreating at such streams.

Algae levels held to 150 mg Chl \textsubscript{a}/m\textsuperscript{2} or less should also better protect the agriculture use. Less filamentous algae means irrigation systems, which are often operating along eutrophied streams and rivers in the summer in Montana, will become clogged much less often. Keeping these irrigation systems clean of algae is inefficient and costly.

The public perception survey did not directly address eastern Montana prairie streams which are quite different in appearance from their western counterparts. Montana prairie streams often become intermittent, are generally low gradient, typically have mud bottoms and are turbid, frequently have substantial macrophyte populations, and support fishes such as bullhead, walleye, chubs, bass and other fish preferring summer temperature 18\degree C or greater\textsuperscript{41,79}. It is not uncommon in these streams to see macrophytes intermixed with filamentous algae and floating masses of green algae; these types of conditions are common even in prairie streams minimally impacted by people (i.e., prairie reference streams). Because prairie streams are fundamentally different in many ways from western Montana trout streams\textsuperscript{41}, the results from the public perception algae survey should probably not be directly applied to them. Prairie streams nevertheless have important and sensitive beneficial uses that need to be protected. Prairie streams have a diverse array of fish and aquatic life that can be harmed by eutrophication. Harm-to-sensitive use thresholds for prairie streams should therefore be defined by those existing water quality criteria for DO, pH, and total dissolved gas (TDG) already adopted as standards\textsuperscript{8} intended to protect fish and aquatic life. Since these effect criteria have been linked to nutrients in prairie streams (e.g., Appendix A), numeric nutrient criteria can be determined.
Section 6.0
Stressor-Response Studies and Reference Site Data as Complementary Components in Determining Numeric Nutrient Criteria

6.1 Stressor-Response Studies

Stressor-response studies examine the relationship between a variable that has the potential to cause a water quality problem (stressor) and the specific effect that it manifests (response). In this work, the stressors of interest are nutrients and the responses are the measurable impacts, i.e. harm, to a stream or river beneficial use. Several stressor-response studies have already been discussed in Section 2.0. When it comes to developing stream and river numeric nutrient criteria, the most useful studies tend to be those that have been carried out in the field (i.e., not in laboratories). Field-based nutrient stressor-response studies vary in the degree of control the researcher has over the study and can be broadly categorized (from most to least controlled) as: artificial stream studies (e.g., 62, 80); whole-stream fertilization experiments (e.g., 13, 14); and “mensurative experiments” 81. Mensurative experiments or studies are those in which the researcher seeks to define a quantitative relationship between an ecological response variable (e.g., stream trout density) and a gradient of an environmental condition (e.g., total P 28). All three of these study types were considered when developing Montana’s numeric nutrient criteria and there are many, many such studies that have been carried out worldwide. For developing Montana’s numeric nutrient criteria, however, we focused on studies that could be used to relate stream nutrient concentrations to beneficial uses, and that have some regional relevance 14, 28, 43, 57, 78, 80-84. The most important of these are shown in Table 6.1. How the studies were used to help with nutrient criteria development will be discussed in Section 6.3.
6.0 Stressor-Response Studies and Reference Site Data as Complementary Components in Determining Numeric Nutrient Criteria

Table 6.1 Studies Addressing Nutrients and Eutrophication that Were Useful for Developing Montana’s Numeric Criteria.

<table>
<thead>
<tr>
<th>Scientific Study</th>
<th>Nutrient(s) Concerned</th>
<th>Study Summary</th>
<th>Type of Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perrin et al. (1987)</td>
<td>N, P</td>
<td>Nitrogen and P are added to a small river in British Columbia, Canada, to observe the effect on benthic algal biomass &amp; algal population structure, as well as salmonid fish production.</td>
<td>Whole-Stream Fertilization Study</td>
</tr>
<tr>
<td>Welch et al. (1989)</td>
<td>P</td>
<td>Results from artificial stream channels are used to help model benthic algal growth in the Skykomish River, Washington. Modifications to the model are made to match with the river’s specific conditions. Emphasis is on soluble P.</td>
<td>Artificial Stream Study, Adapted to a River</td>
</tr>
<tr>
<td>Bolwell, M.L. (1989)</td>
<td>P</td>
<td>Artificial channels built alongside and using water from the S. Thompson River in British Columbia, Canada. The channels are dosed with P to determine the effect on peak areal biomass of benthic cyanobacteria.</td>
<td>Artificial Stream Study</td>
</tr>
<tr>
<td>Watson et al. (1990)</td>
<td>N, P</td>
<td>Water from the Clark Fork River, Montana is pumped at a location where nutrient concentrations are low (due to the influence of two tributaries) into artificial stream channels. The artificial channels are then dosed with soluble N and P and the benthic algal biomass changes are measured.</td>
<td>Artificial Stream Study</td>
</tr>
<tr>
<td>Chiflet et al. (1999)</td>
<td>P</td>
<td>Benthic algal biomass and species composition compared to nutrient concentrations in 13 Canadian rivers. Correlations between algae and nutrients given.</td>
<td>Mesure Study</td>
</tr>
<tr>
<td>Sotoa, A. (2002)</td>
<td>N, P</td>
<td>A 16 year study on the Bow River (Alberta, Canada) quantifying the reduction in biomass of benthic algae and aquatic macrophytes resulting from reduced N and P concentrations in the discharge of two municipal wastewater treatment plants.</td>
<td>Mesure Study</td>
</tr>
<tr>
<td>Dodds et al. (2006)*</td>
<td>N, P</td>
<td>A large database containing benthic algae and N and P data from hundreds of temperate streams worldwide is used to define regression-equation relationships between total N and P concentrations and benthic algae levels.</td>
<td>Mesure Study</td>
</tr>
<tr>
<td>Wang et al. (2007)</td>
<td>N, P</td>
<td>240 wadeable streams in Wisconsin are systematically sampled for N &amp; P, macroinvertebrates, and fish. A series of correlations between the nutrients and the biological assemblages are presented.</td>
<td>Mesure Study</td>
</tr>
<tr>
<td>Suppe et al. (2003) Unpublished data, Appendix A, this document</td>
<td>N</td>
<td>Relationships between diatom algae and stream environmental characteristics (including nutrients) are presented for a group of NE Montana prairie streams. Diatom population characteristics (species) are then used to infer DO concentrations in the streams.</td>
<td>Mesure Study</td>
</tr>
</tbody>
</table>


6.2 Reference Sites

DEQ has been working for nearly 20 years to locate and characterize wadeable streams which have little or no human disturbance. Some work was completed in the early 1990s and involved...
the collection of water quality and biological data at stream sites considered by regional land managers to be minimally disturbed. Little was done through the remainder of the 1990s, but the work was recommenced in 2000 and continues to this day in an updated guise using better defined and much more rigorous screening methods compared to the earlier undertaking. Over 130 reference stream sites have so far been identified around Montana (see Figure 4.1, page 20). Reference sites represent our best approximation of stream condition in the absence of substantial human disturbance or alternation, although they are not all pristine. In the selection of reference sites, human activities are considered an integral part of the landscape as long as those activities do not negatively harm the various uses of the water (drinking, aquatic life, fisheries, recreation, etc.). DEQ assesses each candidate site and those that pass (i.e., are considered final reference sites) are ranked as either tier 1 or 2 in accordance with how well they fit one of the following definitions:

**Tier 1 — Natural Condition:** The characteristics of a waterbody that is unaltered from its natural state, or there are no detectable human-caused changes in the completeness of the structure and function of the biotic community and the associated physical, chemical, and habitat conditions. All numeric water quality standards must be met and all beneficial uses must be fully supported unless impacts are clearly linked to a natural source. The natural condition is the highest attainable biological, chemical, physical, and riparian condition for waterbodies.

**Tier 2 — Minimally Impacted Condition:** The characteristics of a waterbody in which human activities 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 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 waterbody’s best potential condition.

Montana reference sites represent an array of stream sizes, having Strahler stream orders that range from 1st through 6th (Figure 6.1); most are 3rd order. They occur in all of Montana’s level III ecoregions, except for in the Wyoming Basin which has only a small extent in SE Montana (Figure 4.1).
6.0 Stressor-Response Studies and Reference Site Data as Complementary Components in Determining Numeric Nutrient Criteria

6.2.1 The Reference Site Nutrient Database

DEQ has assembled a database containing all nutrient data collected from Montana reference sites. This database has been rigorously screened to assure data quality. It was noted during an early phase of the database’s development that some sites among the network of reference sites contributed — for a variety of reasons — a disproportionate amount of nutrient data to the whole than did others. For example, one site may have been sampled 20 times while other sites were only sampled once or twice. This usually occurred because a few sites had a long history of nutrient sampling, while others had only been identified and sampled in recent years. Equitable representativeness among the sites was important for proper reference characterization of each ecoregion. So, in 2007, reference stream sites were sampled in a targeted manner for a suite of nutrients (TN, TP, TKN, NO₂⁻, SRP, and ammonia) with the intent of making each site a significant contributor to the aggregate nutrient dataset. We used the Brillouin evenness index (J), calculated on an ecoregion-by-ecoregion basis, to measure our success. Very uneven datasets have J values near zero (e.g., 0.2), while a dataset with a J value of 1.0 means each site contributes exactly the same number of samples to the total. Our goal was to achieve index values of ≥ 0.8 (80% evenness) for each level III ecoregion. The work was very successful (Table 6.2), and gives DEQ confidence that the 2008 database has good dispersion of sampling effort among the reference sites and good overall representation of the range of nutrient concentrations found across all reference sites.
6.0 Stressor-Response Studies and Reference Site Data as Complementary Components in Determining Numeric Nutrient Criteria

Table 6.2 Nutrient Sampling Indices for Montana Reference Sites, Before & After 2007 Sampling. Data are Presented by Level III Ecoregion.

<table>
<thead>
<tr>
<th>Ecoregion (Level III)</th>
<th>Season*</th>
<th>Status as of November 2006</th>
<th>Subsequent to 2007 Sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Brilliance Evenness Index (J)</td>
<td>Proportion of Sites Providing Zero Samples</td>
</tr>
<tr>
<td>Middle Rockies</td>
<td>Growing</td>
<td>0.65</td>
<td>14%</td>
</tr>
<tr>
<td>Northern Rockies</td>
<td>Growing</td>
<td>0.62</td>
<td>0%</td>
</tr>
<tr>
<td>Canadian Rockies</td>
<td>Growing</td>
<td>0.38</td>
<td>31%</td>
</tr>
<tr>
<td>Idaho Batholith</td>
<td>Growing</td>
<td>0.72</td>
<td>0%</td>
</tr>
<tr>
<td>NW Glaciated Plains</td>
<td>Growing</td>
<td>0.72</td>
<td>29%</td>
</tr>
<tr>
<td>NW Great Plains</td>
<td>Growing</td>
<td>0.72</td>
<td>36%</td>
</tr>
</tbody>
</table>

6.3 Integrating Information from Stressor-Response Studies and Reference Sites

Stressor-response studies provide information on the effect nutrients have in streams. Stream reference sites confer an understanding of what nutrient concentrations are like in the absence of substantial human disturbance. For the purpose of setting criteria over a large and diverse landscape, however, each of these individual pieces of information is, in a sense, incomplete. Stressor-response studies provide the scientific understanding as to how eutrophication is manifested in streams, but each study has its own statistical uncertainties (e.g., these studies often have only a moderate correlation coefficient, or $r^2$, between nutrient and response variable), they are usually limited in scope (specific to a particular region or individual stream), and there are usually only a few studies available in any given ecoregion. In contrast, reference sites — if there are enough of them — provide good landscape coverage for an array of unimpacted regional stream types, but they do not by themselves tell us about thresholds of harm to beneficial uses. But when these two types of information are brought together, a very powerful tool is created that affords good confidence about stream ecology, eutrophication effects, and when beneficial uses become harmed. This section will demonstrate how reference and stressor-response data were integrated in order to derive Montana’s numeric nutrient criteria.

6.3.1 Nutrient Concentrations from Stressor-Response Studies Compared to Concentrations from Reference Sites (or, Integrating the Stressor-Response and Reference Approaches)

Figure 6.2 is a conceptual diagram showing how nutrient concentrations from reference sites and nutrient concentrations that harm uses (derived from stressor-response studies) might be related to one another. Because reference streams are, by definition, minimally impacted by people and support their beneficial water uses, it is logical that the majority of the nutrient concentration data collected from reference sites would be acceptable. Intuitively, nutrient concentrations that harm stream uses should only be among the highest concentrations observed at reference sites or, perhaps, be even higher than any concentration observed in reference sites (Figure 6.2). By comparing harm-to-use nutrient concentrations derived from stressor-response studies to applicable reference site distributions, one can gain more confidence that the stressor-response study values are correct. For example, if a particular stressor-response study suggests that a total

November, 2008  34
P concentration of 1.0 µg/L is needed to protect beneficial stream uses, but that concentration falls at the 25<sup>th</sup> percentile of the applicable distribution of reference sites (i.e., on the far left side of the histogram in Figure 6.2), it would lead one to suspect that either there was a problem with the stressor-response study results or the quality of the reference sites. In this manner, stressor-response results and reference data are used to complement and cross-check one another.

Figure 6.2 Conceptual Diagram Showing a Nutrient Concentration Histogram for Reference Sites. The figure shows where along the x-axis, relative to the reference-data histogram, nutrient concentrations likely to harm beneficial water uses would be expected to be found.

The cross-comparison analysis described above was undertaken several years ago for Montana and showed that, on average, nutrient concentrations at the 86<sup>th</sup> percentile of the reference distribution were equivalent to harm-to-beneficial use thresholds<sup>60</sup>. Referring again to Figure 6.2, empirical data analysis indicates that among reference sites harm-to-use nutrient concentrations are not greater than (beyond) the aggregate reference-site distribution, but instead are among the very highest concentrations measured in reference sites. It is not surprising that some reference sites have some nutrient samples whose concentrations are higher than the harm-to-use threshold identified using stressor-response studies. In any population of data there are always low and high values that differ considerably from the population’s central tendency; the important point is that nutrient concentrations in reference sites that are greater than the harm-to-use threshold occur infrequently, e.g. due to an atypical high-flow event in summer. It is when the harm-to-use concentrations occur commonly in a stream that eutrophication problems occur (e.g., see Section 4.2.3 in Appendix H of Varghese and Cleland<sup>61</sup>).

Since the work described above<sup>60</sup> was published, DEQ has made improvements to the reference site nutrient database (see Section 6.2.1), there has been a new stressor-response study completed...
6.0 Stressor-Response Studies and Reference Site Data as Complementary Components in Determining Numeric Nutrient Criteria

(Appendix A), and our understanding of harm-to-use thresholds is improved due to the nuisance algae public perception study77; these changes warrant another iteration of the analysis. Table 6.3 shows key descriptive statistics for the current reference-site data (2008 database). Table 6.4 shows the relationship between regional stressor-response studies and regional reference datasets. In Table 6.4, for each stressor-response study, the harm-to-use nutrient concentration threshold was derived as (1) that concentration that would maintain benthic algae levels ≤ 150 mg Chl a/m² or (2) that nutrient concentration that would maintain DO concentrations at state standards. Each of the stressor-response studies shown was carried out in and specific to an ecoregion that occurs in Montana. The nutrient concentration at the harm-to-use threshold from each study was matched to the equivalent concentration in its corresponding reference-site nutrient distribution. By “corresponding” reference site nutrient distribution, we mean all data in our database from reference sites located in the same ecoregion where the stressor-response study took place, collected during the same time of year (growing season, i.e., summertime).

Table 6.3 Descriptive Statistics for Nutrients from Montana Reference Sites, 2008 Database. Data are Presented by Level III Ecoregion.

<table>
<thead>
<tr>
<th>Descriptive Statistics, by Nutrient</th>
<th>Level III Ecoregion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Northern Rockies</td>
</tr>
<tr>
<td>TOTAL P</td>
<td></td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>88</td>
</tr>
<tr>
<td>minimum (mg/L)</td>
<td>0.001</td>
</tr>
<tr>
<td>25th percentile (mg/L)</td>
<td>0.003</td>
</tr>
<tr>
<td>50th percentile (mg/L)</td>
<td>0.003</td>
</tr>
<tr>
<td>75th percentile (mg/L)</td>
<td>0.007</td>
</tr>
<tr>
<td>maximum (mg/L)</td>
<td>0.018</td>
</tr>
<tr>
<td>TOTAL N</td>
<td></td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>38</td>
</tr>
<tr>
<td>minimum (mg/L)</td>
<td>0.012</td>
</tr>
<tr>
<td>25th percentile (mg/L)</td>
<td>0.031</td>
</tr>
<tr>
<td>50th percentile (mg/L)</td>
<td>0.050</td>
</tr>
<tr>
<td>75th percentile (mg/L)</td>
<td>0.116</td>
</tr>
<tr>
<td>maximum (mg/L)</td>
<td>0.360</td>
</tr>
<tr>
<td>NO$_2$+3</td>
<td></td>
</tr>
<tr>
<td><strong>n</strong></td>
<td>89</td>
</tr>
<tr>
<td>minimum (mg/L)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>25th percentile (mg/L)</td>
<td>0.005</td>
</tr>
<tr>
<td>50th percentile (mg/L)</td>
<td>0.009</td>
</tr>
<tr>
<td>75th percentile (mg/L)</td>
<td>0.011</td>
</tr>
<tr>
<td>maximum (mg/L)</td>
<td>0.150</td>
</tr>
</tbody>
</table>

Note: n represents all samples from all reference sites combined together.
### Table 6.4 Regional Stressor-Response Study Nutrient Concentrations and their Corresponding Percentile Values in Corresponding Reference-Site Nutrient Frequency Distributions.

<table>
<thead>
<tr>
<th>Stressor-response Study</th>
<th>Nutrient</th>
<th>Notes on Study</th>
<th>Stressor-response Study Nutrient Concentration (μg/L)</th>
<th>Season of Application</th>
<th>Level III Ecoregion</th>
<th># Samples in Distribution During Growing Season</th>
<th>Percentile in Reference Distribution Matching</th>
<th>Beneficial Use the Nutrient Concentration Threshold Applies To</th>
</tr>
</thead>
<tbody>
<tr>
<td>Welch et al. (1989)</td>
<td>SRP</td>
<td>The SRP concentration would constrain the distance the Spokane River has algal biomass of 150 mg Chl a/m² to about 16 km.</td>
<td>0.01</td>
<td>Growing</td>
<td>Northern Rockies</td>
<td>75</td>
<td>24&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Recreation</td>
</tr>
<tr>
<td>Watson et al. (1990)</td>
<td>SRP</td>
<td>The SRP concentration corresponding to algal standing crop of 150 mg Chl a/m²</td>
<td>0.011</td>
<td>Growing</td>
<td>Middle Rockies</td>
<td>21</td>
<td>97&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Recreation</td>
</tr>
<tr>
<td>Sondak, A. (2002)</td>
<td>TP</td>
<td>Based on a nutrient vs. benthic-algae regression equation, TP concentration would maintain algal standing crop ≤ 150 mg Chl a/m² on the Bow River near Calgary, Alberta, Canada.</td>
<td>0.018</td>
<td>Growing</td>
<td>Canadian Rockies</td>
<td>68</td>
<td>97&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Recreation</td>
</tr>
<tr>
<td>Supplee, M.W. (2008)</td>
<td>TN</td>
<td>TN concentration would prevent dissolved oxygen (DO) from dropping below state standards in prairie streams. Quantitative relationships (correlation, changepoint analysis) between diatom-inferred DO and TN concentrations were used to derive the TN concentration.</td>
<td>1.12</td>
<td>Growing</td>
<td>Northwestern Plains</td>
<td>59</td>
<td>70&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Fish &amp; Aquatic Life</td>
</tr>
</tbody>
</table>

* See Table 4.1, this report, for the start and end dates of the Growing Season.

1 The Bow River at Calgary is downstream of the boundary of the level III ecoregion 'Canadian Rockies'. Stressor-response study TP recommendation was assigned to the Canadian Rockies ecoregion since the majority of the river's drainage area upstream of Calgary is in the Canadian Rockies.
The results in Table 6.4 show that harm-to-use-threshold nutrient concentrations equal concentrations at the 87th (mean) and 91st (median) percentile of reference. If only the mountainous ecoregions are considered (Northern, Canadian, and Middle Rockies), harm-to-use-threshold nutrient concentrations correspond to the 93rd (mean) and 94th (median) percentile of reference, with a low CV of 5.5%. The single study from low-gradient, warm-water prairie streams (Appendix A) has a notably lower reference percentile match (70th of reference) compared to the other, mountainous studies (87th-97th of reference)(Table 6.4). This may have resulted because (1) the prairie stream study is looking at a different cause of harm (minimum DO vs. nuisance benthic algae levels) than the other studies, (2) the empirical relationship between reference site nutrient data and stressor-response derived nutrient concentrations is inherently different in prairie streams, (3) all the reference sites in the prairie ecoregions fit the Tier 2 definition (some human impacts noted) whereas in the mountain ecoregions there is a mix of Tier 1 and Tier 2 sites, or (4) this is an unusually low reference-to-stressor response match, something of an outlier, and other studies we might carry out in prairie streams in the future may show results more like the other studies in Table 6.4.

Overall, nutrient concentrations suggested by the individual stressor-response studies become more convincing when they are viewed through their collective relationship to reference. It is very unlikely that the pattern seen in Table 6.4, wherein nutrient concentrations from four regionally-applicable stressor-response studies are clustered among the upper percentiles of four different reference distributions from four different ecoregions, would occur by chance\(^6\). Thus, it can be reasonably concluded that (1) stressor-response derived harm-to-use nutrient concentrations are consistently found among the upper percentiles of applicable reference site distributions\(^6\), and (2) concentrations in the upper percentiles of reference site nutrient distributions can act as surrogates for harm-to-use concentrations. On a statewide basis, the data suggest the 90th percentile of reference (midrange between the mean and median of the case study-to-reference matches) is a good starting point for determining appropriate statewide criteria.

### 6.3.2 Other Information Reviewed to Cross-check the Criteria

The criteria were cross-checked using the Redfield ratio\(^8\),\(^9\),\(^0\), which has long been used to assess which nutrient is likely to limit algal growth. Optimal nutrient ratios (by weight) in benthic stream algae are about 54:8:1 (carbon:nitrogen:phosphorus)\(^9\),\(^1\),\(^2\), similar to the widely-accepted phytoplankton algal ratio of 47:7:1. “Optimal” means that, from an alga’s perspective, all three macronutrients are sufficiently available in the environment to allow maximum growth (i.e., none of the elements is in short supply). Although benthic stream algae have optimal Redfield N:P ratios of about 8, N:P ratios ranging from 6 to 10 mean neither element is strongly limiting\(^9\).

\(^e\) A simple way to calculate the probability of this outcome is by considering each case study as an independent event. If it is assumed that the probability that the nutrient concentration derived from any given stressor-response study has, by random chance, a 0.75 probability of falling somewhere between the 1st and 75th percentile of the corresponding reference distribution, and a 0.25 probability of falling between the 76th and 100th percentile, then the likelihood of the pattern seen in Table 6.4 becomes \(0.25 \cdot 0.25 \cdot 0.25 \cdot 0.75 = 0.01\) (1% chance).
Ideally, nutrient criteria should be set so that N:P concentration ratios are near optimal or, if one nutrient tends to be limiting in a region, the criteria should lean towards controlling the limiting nutrient. To accomplish the latter, the N:P ratio should generally be higher than Redfield ratio to control P and lower than Redfield to control N. We examined the N:P ratio of nutrient concentrations at the 90th percentile of reference to see if we would maintain appropriate nutrient ratios, relative to the Redfield ratio. We constrained this analysis to the four mountainous level III ecoregions (Northern, Canadian, and Middle Rockies, and the Idaho Batholith) where control of nuisance benthic algae is the goal. In these ecoregions, TN: TP ratios of the concentrations at the 90th percentile of reference were either within the optimal range (between 6 and 10) or were high, meaning the criteria would tend to control algae via P limitation. The high ratios occurred in the Northern Rockies, Canadian Rockies, and Idaho Batholith, where the TN:TP ratios ranged from 12-35. For these ecoregions this situation is acceptable, even desirable, as several regional studies show that P limitation is common here and benthic algae respond quickly to small increases in P. Another check of the criteria was made by assessing whether or not total P concentrations at the 90th of reference would maintain soluble P concentrations low enough to assure an effect on algae. Studies, many carried out in this region, show soluble reactive phosphate (SRP) should be kept under 5 µg/L, no greater than 11 µg/L, or perhaps as high as 22 µg/L (if grazers are present) in order to maintain benthic algae levels (including Cladophora) at or below 150 mg Chl a/m² of stream bottom. We again focused on the four western Montana ecoregions (Northern, Canadian, and Middle Rockies, and Idaho Batholith) as their streams most resemble those in the cited studies. SRP:TP ratios in rivers and streams worldwide range from about 0.1 to 0.7. In Montana, long-term monitoring at river and stream sites show SRP:TP ratios typically range from 0.26 to 0.5. We used an SRP:TP ratio of 0.35, which we considered to be a good regional average. Multiplying 0.35 by the ecoregional TP concentration at the 90th percentile of reference (during the growing season, i.e., summertime) resulted in calculated SRP concentrations of: 2 µg/L (Canadian Rockies); 4 µg/L (Idaho Batholith); 4 µg/L (Northern Rockies), and 17 µg/L (Middle Rockies). All fell below the more conservative SRP benchmark (5 µg SRP/L), except the Middle Rockies, which was much higher. Because it seemed elevated, the SRP concentration (17 µg/L) calculated for the Middle Rockies was further evaluated relative to a highly-applicable artificial stream study (the study was carried out in the Middle Rockies ecoregion in Montana). That study shows 17 µg SRP/L might still keep algae below 150 mg Chl a/m², as 17 µg SRP/L falls within the 95% confidence interval of the study’s benthic chlorophyll measurements. We also reviewed studies that did not occur specifically in a Montana ecoregion but were carried out in northern temperate rivers and streams and provide good comparative information. We compared the TN and TP concentrations at the 90th percentile of reference for the Middle Rockies ecoregion to results from other temperate-stream studies (Table 6.5). These studies occurred streams roughly comparable to those found in the Middle Rockies. (The Middle Rockies ecoregion is the largest ecoregion in western Montana and has more land area than the Northern Rockies, Canadian Rockies, and Idaho Batholith combined.)
6.0 Stressor-Response Studies and Reference Site Data as Complementary Components in Determining Numeric Nutrient Criteria

Table 6.5  Nutrient Concentrations from Studies Carried out in Northern Temperate Rivers & Streams Compared to Nutrient Concentrations at the 90th Percentile of Reference Sites from Montana’s Proportion of the Middle Rockies Ecoregion.

<table>
<thead>
<tr>
<th>Study</th>
<th>Where Study Took Place</th>
<th>Notes on Study</th>
<th>Nutrient (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>This Document</td>
<td>Middle Rockies Ecoregion, Montana</td>
<td>90th percentile of reference</td>
<td>0.320 0.048</td>
</tr>
<tr>
<td>Perrin et al. (1987)</td>
<td>British Columbia, Canada</td>
<td>Total N and total P concentrations quantitatively added to a small, low-nutrient river and resulted in peak benthic algae of 150 mg Chl a/m² and a shift towards dominance by filamentous algae.</td>
<td>0.4 0.02</td>
</tr>
<tr>
<td>Müller &amp; Rankin (1998)</td>
<td>Ohio</td>
<td>Nutrient concentration threshold beyond which deleterious effects on fish communities are observed.</td>
<td>n/a 0.06</td>
</tr>
<tr>
<td>Chételar et al. (1999)</td>
<td>Ontario &amp; Quebec, Canada</td>
<td>Benthic algal biomass and nutrient concentrations examined in 13 rivers. Moderately strong relationship (r² = 0.56) found between total P and benthic Chl a levels. TP concentration shown would maintain algae at 150 mg Chl a/m².</td>
<td>n/a 0.07</td>
</tr>
<tr>
<td>Dodds et al. (2006)</td>
<td>Data from North American, Australian, New Zealand and European temperate streams</td>
<td>Based on a nutrient vs. benthic-algae regression equation, TN and TP concentrations would maintain benthic algae ≤ 150 mg Chl a/m² (maximum). Concentrations determined using the 90th listed equation of the literature dataset.</td>
<td>0.578 0.080</td>
</tr>
<tr>
<td>Wang et al. (2007)</td>
<td>Wisconsin</td>
<td>Total N and total P concentration thresholds where the largest change in biometrics occur and beyond which fish and macroinvertebrate assemblages are likely to be degraded.</td>
<td>0.99 0.073</td>
</tr>
</tbody>
</table>

From each of the studies in Table 6.5, nutrient concentrations that represent a biological impact threshold for fish or macroinvertebrates were used, or, alternatively, nutrient concentrations that would result in benthic algae of no more than 150 mg Chl a/m² (see “Notes on Study” in the table). Overall, the Middle Rockies TP concentration at the 90th of reference (0.048 mg/L) falls within the range of the other studies, while the TN concentration (0.32 mg/L) is at the low end. In all cases, total N or total P concentrations are within the same order of magnitude.

Comparable studies specific to prairie streams are few. The study in Appendix A (this report) suggests 1.12 mg TN/L as a threshold concentration that should prevent nighttime DO minima from dropping below state standards. Other prairie stream studies recommend similar values. One study analyzes a subset of the data in Appendix A (from the Milk/Lower Missouri Basin), and also data from the Sheyenne River basin in North Dakota, and suggests TN should be held to about 1 mg/L. The 1 mg TN/L threshold was determined based on (1) a sharp decline, at 1.03 mg TN/L, in a correlation between TN and pollution-sensitive diatoms from the Milk/Lower Missouri basin, and (2) due to the decline in macroinvertebrate EPT taxa with increasing TN concentrations in the Sheyenne River basin. Another study uses a weight-of-evidence approach, compiling literature values, median reference stream concentrations, etc., and recommends 0.96 mg TN/L as the concentration most likely to protect use and integrity of prairie streams99.
Finally, we considered other states’ approaches. The Tennessee Department of Environment and Conservation finds that nutrient concentrations at the 90th percentile of stream reference sites correspond well to harm-to-use thresholds for their wadeable streams. Like DEQ, they also stratify their regional nutrient expectations using ecoregions. Their work shows that once nutrient concentrations exceed the 90th of reference, streams generally show aquatic life impairment based on their macroinvertebrate biointegrity metrics.

From this series of comparisons, it can be reasonably concluded that N and P concentrations found among the upper percentiles of reference site nutrient distributions can act as surrogates for harm thresholds of sensitive beneficial uses. As discussed in Section 6.3.1, comparing stressor-response data to their corresponding reference data adds strength to the conclusions drawn from the individual stressor-response studies. Further, using concentrations that are linked to the reference distribution assures that localized, regional landscape effects on background nutrient concentrations will be reflected in the criteria, assuring that the criteria will not be overly stringent or insufficiently protective.
Section 7.0
Criteria Specifications for Montana’s Ecoregions

Previous sections detailed how we identified a rational geospatial frame for segregating nutrient concentrations (Section 4.0), how we identified eutrophication’s harm to beneficial uses (Section 5.0), and how we zeroed in on appropriate nutrient criteria concentrations (Section 6.0). This section will specify criteria expectations for different regions of Montana, and includes recommended effect variables (e.g., nuisance algae limits) that should accompany the nutrient criteria in some regions.

Table 7.1 below shows the criteria recommendations for wadeable streams of the state, including one example from a level IV ecoregion. Note that benthic algae criteria are also provided; these should accompany the nutrient criteria to assure use protection. The details of how these values were arrived at are discussed in the following sub-sections. As for all water quality criteria, numeric nutrient criteria will undergo periodic revision and update as more stressor-response studies are completed and more reference data is collected.

<table>
<thead>
<tr>
<th>Level III Ecoregions</th>
<th>Nutrient Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ecoregion</strong></td>
<td>Reference Period When Criteria Apply</td>
</tr>
<tr>
<td>Northern Rockies</td>
<td>July 1 - Sept. 30</td>
</tr>
<tr>
<td>Canadian Rockies</td>
<td>July 1 - Sept. 30</td>
</tr>
<tr>
<td>Middle Rockies</td>
<td>July 1 - Sept. 30</td>
</tr>
<tr>
<td>Idaho Batholith</td>
<td>July 1 - Sept. 30</td>
</tr>
<tr>
<td>Northwestern Glaciated Plains</td>
<td>June 16 - Sept. 30</td>
</tr>
<tr>
<td>Northwestern Great Plains, Wyoming Basin</td>
<td>July 1 - Sept. 30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level IV Ecoregions</th>
<th>Nutrient Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-calcareous Foothill Grassland (Parent Level III: Northwestern Great Plains)</td>
<td>July 1 - Sept. 30</td>
</tr>
</tbody>
</table>
7.1 Streams in Montana’s Mountainous Level III Ecoregions

Wadeable streams in the largely mountainous level-III ecoregions Canadian Rockies, Northern Rockies, Middle Rockies, and Idaho Batholith have recreation and fisheries among their beneficial uses. They generally support salmonid fishes, commonly have moderate gradient and gravel bottoms (Figure 7.1), and are used for recreation of all kinds (fishing, swimming, boating, etc.). Preventing harm to the recreation use and the fish and aquatic life use in this region is very important. In Section 6.0 we presented a series of arguments that support the idea that nutrient concentrations at the 90th percentile of reference are, overall, equivalent to harm-to-use thresholds. The 90th percentile is also near the low end of the error bar (CV; ± 5.5%) around the mean reference-to-stressor response match (93rd percentile) specific to the four mountainous ecoregions in Montana (see page 38). Therefore, we recommend that nutrient concentrations in the Canadian Rockies, Northern Rockies, Middle Rockies, and Idaho Batholith be set at the 90th percentile of each ecoregion’s reference dataset (see Table 7.1). Nutrient concentrations held to the 90th of reference should assure that benthic algae levels do not exceed 150 mg Chl a/m². In addition to the TP, TN, and NO₂+³ criteria, a benthic algae criterion of 150 mg Chl a/m² should be adopted for these streams to prevent nuisance growth, per DEQ’s algae perception survey⁷⁷. Benthic algae should be measured and reported using DEQ Standard Operating Procedures. All criteria (TP, TN, NO₂+³, and benthic algae levels) should apply during the “Growing Season” of each ecoregion (Table 4.1), as that is when the use-impact (nuisance benthic algae) is typically manifested.

Figure 7.1 Willow Creek. A reference stream site in the level III ecoregion Middle Rockies.
7.0 Criteria Specifications for Montana’s Ecoregions

7.2 Prairie Streams in Eastern Montana Level III Ecoregions

Wadeable prairie streams are found in eastern Montana’s level III ecoregions Northwestern Glaciated Plains and Northwestern Great Plains. Due to its very small size in Montana (Figure 4.1), limited data, and lack of reference sites, the level III ecoregion Wyoming Basin should receive the same criteria as the adjacent Northwestern Great Plains ecoregion. Streams in this overarching region have recreation and fisheries among their beneficial uses. Many become intermittent, are generally low gradient, typically have mud bottoms and are turbid (Figure 7.2), and often have substantial macrophyte populations. The streams are mostly classified as marginal- or non-salmonid fisheries (e.g., C-3 classification; ARM 17.30.629[1]) meaning, in general, they should support warm-water fish and associated aquatic life. It is not uncommon in these streams to see macrophytes intermixed with filamentous algae and floating masses of green algae; these types of conditions are common even in prairie reference streams. Because prairie streams are fundamentally different in many ways from western Montana trout streams, the results from the public perception algae survey should probably not be directly applied to them.

The single, applicable stressor-response study we have (Appendix A) shows that a harm-to-use threshold, based on assuring that DO concentrations remain above state minima in order to protect fish and aquatic life, occurs at the 70th percentile of the regional reference TN distribution (Table 6.4). Because the nutrient concentration derived from this stressor-response study has a lower match to reference (70th) compared to the other studies (87th-97th), nutrient concentrations at the 90th of reference are almost certainly too high to protect this region’s beneficial uses. For example, TN at the 70th percentile of reference in the Northwestern Glaciated Plains equals 1.12 mg/L, but the TN concentration at the 90th of reference is 1.91 mg/L — in all likelihood much too high to protect fish and aquatic life. Concentrations at the 75th percentile of reference, on the other hand, appear to be appropriate for prairie streams, for three reasons. First, EPA recommends the 75th percentile of reference as generally appropriate for setting numeric nutrient criteria. Second, the TN concentration at the 75th percentile of reference for the Northwestern Glaciated Plains (1.31 mg/L) falls within the 90% confidence interval (0.78-1.48 mg TN/L) around the harm-to-use threshold (1.12 mg TN/L) identified using changepoint analysis in DEQ’s study (Appendix A). Third, if one considers the overall statewide pattern, it is clear that other stressor-response studies applicable specifically to Montana indicate higher percentile matches to reference (87th-97th).

Given EPA’s general recommendation (the 75th), the statistical uncertainties associated with a single regional study, and the overall statewide stressor response-to-reference pattern, it seems reasonable to use for the prairie streams a somewhat higher value than the 70th. We concluded that nutrient concentrations at the 75th of reference (e.g., TN = 1.311 mg/L, Northwestern Glaciated Plains) should protect fish and aquatic life uses in eastern Montana’s prairie streams. The nutrient criteria (TN, TP, and NO2+3) should apply during the Growing Season (Table 4.1) when aquatic plant growth is heaviest and is most likely to impact DO concentrations. One more consideration: the other prairie stream studies we reviewed (Section 6.3.2), few that there are, suggest TN should be held to around 1 mg/L, which is lower than our final recommendation. Thus, we also recommend that another prairie stream stressor-response study be carried out in this region.

November, 2008
7.3 Mountain-to-Prairie Transitional Streams

Streams that originate in the mountains and flow out to the eastern prairies of Montana (i.e., east of the Rocky Mountain Front) present a special situation for criteria setting. Some streams originate in mountainous level III ecoregions (e.g., Middle Rockies) and then immediately cross into either the Northwestern Glaciated Plains or Northwestern Great Plains (each level III), but the ecology and water quality of the streams within a mountain-to-prairie transitional area still largely reflect mountain-like conditions (Figures 7.3, 7.4). With distance from their mountain source, the streams gradually become more prairie-like. Level III ecoregions are too coarse to capture the transitional nature of these streams, however specific level IV ecoregions do. Level IV ecoregions that contain mountain-to-prairie transitional stream reaches are: Pryor-Bighorn Foothills; Limy Foothill Grassland; Rocky Mountain Front Foothill Potholes; Non-calcareous Foothill Grassland; and Foothill Grassland\textsuperscript{64} (Figure 7.5). Because transitional stream reaches in the level IV ecoregions listed above are highly mountain influenced, criteria should reflect their mountain-like qualities. The algae criterion of 150 mg Chl $a/m^2$ should apply to streams within these specific level IV ecoregions. At present, nutrient data for these regions are scarce and reference sites are few. DEQ continues to target these regions as a high priority for reference site identification and nutrient sampling. The determination of specific nutrient criteria for these and other level IV ecoregions are discussed next.
Figure 7.3  Rock Creek, a Reference Stream Site in the Transitional Level IV Ecoregion Non-calcareous Foothill Grassland (Part of the Northwestern Great Plains).

Figure 7.4  Sweet Grass Creek, a Reference Stream Site in the Transitional Level IV Ecoregion Non-calcareous Foothill Grassland (Part of the Northwestern Great Plains).
7.0 Criteria Specifications for Montana’s Ecoregions

7.4 Identifying Criteria for Unique Level IV Ecoregions

As discussed in Section 4.0, level IV ecoregions have nutrient concentrations that are often demonstrably different from the larger level III ecoregion of which they are a part. When warranted, unique level IV ecoregions will be separated from their level III parent ecoregion. Nutrient concentrations in level IV ecoregions should be set at the 75th or 90th percentile of their particular reference sites depending upon whether the parent level III ecoregion uses the 75th or 90th (see Sections 7.1 and 7.2), except for mountain-to-prairie transitional ecoregions (Section 7.3). For the latter, we recommend the 80th of reference as it accounts for these ecoregions’ transitional nature between mountain (where the 90th percentile is used) and prairie (where the 75th percentile is used). Criteria for all level IV ecoregions should apply during the Growing Season, using the same dates shown in Table 4.1 for their parent level III ecoregion.

Segregating nutrient criteria at the level IV scale will be done carefully. The much smaller spatial scale of level IV ecoregions leads to issues of small sample size and limited statistical power and, in addition, the stressor-response studies presented in Table 6.4 all occurred at the level III scale. In order for nutrient criteria to be broken out at the level IV scale, the following conditions must be satisfied:

- Within the level IV ecoregion there are at least 12 samples for each nutrient group (12 TP samples, 12 TN samples, 12 NO$_{2+3}$ samples). For details on this sample minimum, see Appendix H in Varghese and Cleland.
7.0 Criteria Specifications for Montana’s Ecoregions

- There is a statistically significant difference (90% confidence level) between the level IV ecoregion reference-site nutrient data and those of the parent level III ecoregion. (This analysis has already been completed.)

- The manner in which the level IV nutrient concentrations are different (higher, lower) compared to the parent level III ecoregion makes sense, given the geology, vegetation, soils, climate, and hydrology of the level IV ecoregion.

- If in doubt, e.g. the nutrient concentration for a particular level IV ecoregion doesn’t make sense relative to what would be expected, lump rather than split (i.e., continue to use the level III concentrations rather than split out the level IVs).

- If the level IV ecoregion in question is a mountain-to-prairie transitional one (see list in Section 7.3), the nutrient criteria should be established at the 80th percentile of reference.

An example of a level IV ecoregion meeting the conditions above is the Non-calcareous Foothill Grassland. It is a sub-ecoregion of the Northwestern Great Plains ecoregion. During the growing season (June 16th – September 30th), there are 16 TP samples from reference sites whose concentration at the 80th percentile of reference is 0.04 mg/L (Table 7.1). TP concentrations in the Non-calcareous Foothill Grassland are significantly lower than those of the Northwestern Great Plains, which makes sense given this mountain-to-prairie transitional ecoregion’s position downstream of the mountains.

In the near future, DEQ will assemble a water quality circular containing all the specific nutrient concentrations and their associated regions and times of application (level III or IV ecoregion, season of application, etc.). In assembling the circular, the level IV ecoregion screening process bulleted above will be carried out. Only level IV ecoregions that meet the specified conditions will have level IV ecoregion criteria in the circular. Note also that not all level IV ecoregions have yet been subjected to the full evaluation process outlined above, therefore only one level IV example was provided in Table 7.1.
Section 8.0  
Review and Recommendations

8.1  Review

Beneficial uses are valuable characteristics of a stream or river resource that, directly or indirectly, contribute to human welfare. Examples of beneficial uses include water supply, fish and aquatic life, and recreation. Beneficial uses are established in law and reflect the societal values embodied in those laws. The intent of water quality criteria, in turn, is to assure a level of water quality that will protect the beneficial uses. Some beneficial uses are more sensitive to harm than others; water quality criteria are required by law to protect the most sensitive use from harm.

DEQ has been carefully developing the nutrient criteria discussed in this document over the past eight years; they are applicable to wadeable streams. They do not apply to large rivers, lakes, or wetlands. The criteria are intended to protect wadeable streams from the detrimental and undesirable effects of eutrophication (nutrient enrichment by nitrogen and phosphorus compounds). Nitrogen and phosphorus concentrations in wadeable streams vary a great deal in accordance with regional geology, soils, climate, and vegetation. To address this fact, DEQ has developed and tested a mapping system that will assure that appropriate nutrient criteria are established in different parts of the state only after taking into consideration natural regional landscape variation.

Fundamentally, the nutrient criteria are based on stressor-response scientific studies in which harm to a beneficial use is shown. All regionally applicable stressor-response studies (nitrogen or phosphorus as stressor, beneficial use impact as response) that could be located were reviewed. The results of these studies were then compared to corresponding Montana reference sites to assure their results made sense (i.e., the concentrations derived from the studies did not seem unrealistic given the nutrient concentrations observed in the reference sites). Further, we were able to establish specific linkages between the stressor-response results and the reference site nutrient-data distributions; these linkages help overcome statistical uncertainties in any given stressor-response study, and assure that natural, regional effects on background nutrient concentrations are reflected in the criteria.

In some parts of the state, mainly in the west, the most sensitive beneficial use is recreation. Public opinion analysis shows that the public majority does not want to see excessive algae growth in the gravel-bottomed, clear running, trout-fishery streams common in western Montana. For these types of streams, nutrient criteria were developed that should prevent nuisance algal levels (as defined by the Montana public perception study) from developing and should, therefore, protect the recreation use. The criteria will also protect the fishery, which typically comprises trout, char, and whitefish, from the negative effects of excessive nutrient enrichment (e.g., low dissolved oxygen concentrations). The criteria should also better protect the agricultural use by reducing elevated algae levels that clog irrigation systems.

In other parts of the state, low gradient prairie streams are common. Wadeable prairie streams in Montana often become intermittent, commonly have mud bottoms, are turbid, frequently have
substantial macrophyte populations, usually have filamentous algae but sometimes have only phytoplankton algae, and support catfish, bullhead, walleye, chubs, bass, and other warm water fishes. Because prairie streams are fundamentally different in many ways from western Montana trout streams, the results from the public perception algae survey should probably not be directly applied to them. Prairie streams nevertheless have important and sensitive beneficial uses that need protection, like the diverse species of fish mentioned above. For these types of streams, therefore, the nutrient criteria have been set so that they will maintain dissolved oxygen concentrations that will protect regional fish and aquatic life. These dissolved oxygen concentrations are already established in state law\(^8\). At this time the most sensitive use in prairie streams is considered to be fish and aquatic life, and the nutrient criteria are set to protect them. But because studies in prairie streams are so few, future prairie stream stressor-response studies (which should be carried out) may lead to revised criteria which could be based on sensitive-use endpoints other than dissolved oxygen.

Some wadeable streams are transitional between the mountainous region, which is found mainly in the western part of the state, and the prairie region, which is found in the eastern part of the state. The mapping system mentioned above accounts for these transitional streams and DEQ will, where warranted, set criteria for this specific group. These transitional streams typically have characteristics in common with mountain streams and therefore nuisance algae levels (as determined by the public perception study) should be controlled in them as well. Thus, the most sensitive use for these transitional streams is recreation.

### 8.2 Recommendations

- Omernik ecoregions\(^63\) should be used as the basis for applying the criteria across Montana. Ecoregions at the level III scale should at this time be used as the principal means of applying the criteria. However, level IV (fine-scale) ecoregions should be used where warranted. This is particularly true for streams in the mountain-to-prairie transitional areas of the state. Before level IV ecoregions are selected for application of specific nutrient criteria, they should be subjected to the series of screening evaluations in Section 7.4 to assure that the segregation is warranted.

- Criteria should be established for total nitrogen (TN), total phosphorus (TP), and nitrate + nitrite (NO\(_{2-3}\)). In some ecoregions (detailed below), stream-bottom algae levels should also be set as criteria based on the results of the benthic algae public-perception survey\(^77\). Using appropriate statistical evaluation methods and sufficiently-sized datasets (minimum of 12 for each nutrient is suggested), compliance with nutrient criteria should be undertaken using a 20% allowable exceedence rate. Details on these statistical assessment methods are provided in appendix H and I of another document\(^61\). Stream benthic algae levels should be sampled and analyzed using DEQ Standard Operating Procedures.

- Nutrient criteria should be established as a function of the regionally applicable reference streams. Analysis shows that nutrient concentrations among the upper percentiles of reference-stream nutrient frequency distributions correspond to concentrations that scientific studies show impact water quality and beneficial uses. The advantage of
8.0 Review and Recommendations

linking the stressor-response derived concentrations to the regional reference distribution is that it helps overcome statistical uncertainties in any given stressor-response study, and assures that inherent, regional landscape effects on background nutrient concentrations are reflected in the criteria. This helps to assure that the criteria are not overly stringent or insufficiently protective.

- Across all parts of the state, the criteria (nutrients and benthic algae levels) should apply during the Growing Season (i.e., generally during the summer). The start and end dates of the Growing Season vary somewhat by ecoregion; see Table 4.1.

- Streams in the mountainous level III ecoregions (Northern Rockies, Canadian Rockies, Middle Rockies, and Idaho Batholith) should have TN, TP, and NO$_{2-3}$ criteria established at the 90$^{th}$ percentile of each ecoregion’s reference stream nutrient-concentration distribution. Benthic (i.e., stream-bottom) algae level criteria should be set at $\leq$ 150 mg Chl $a$/m$^2$ ($\leq$ 36 g AFDW/m$^2$). Level IV ecoregions that are separated out from any of the mountainous level III ecoregions listed here should also have the same benthic algae criteria, and nutrient criteria set at the 90$^{th}$ percentile of the level IV ecoregion’s reference-stream nutrient distribution.

- Streams in the level III prairie ecoregions (Northwestern Glaciated Plains, Northwestern Great Plains, and Wyoming Basin) should have TN, TP, and NO$_{2-3}$ criteria established as the 75$^{th}$ percentile of each ecoregion’s reference nutrient-concentration distribution. The Wyoming Basin has only a tiny extent in Montana, and has no reference sites, therefore it should have the same criteria as the adjacent Northwestern Great Plains. The criteria should maintain dissolved oxygen levels at state standards. Level IV ecoregions that are separated out from these level III ecoregions should also have nutrient criteria concentrations set at the 75$^{th}$ percentile of the applicable level IV ecoregion’s nutrient reference distribution, unless they are a mountain-to-prairie transitional ecoregion (discussed in the next bullet).

- Mountain-to-prairie transitional streams found in the level IV ecoregions Pryor-Bighorn Foothills, Limy Foothill Grassland, Rocky Mountain Front Foothill Potholes, Non-calcareous Foothill Grassland, and Foothill Grassland should have TN, TP, and NO$_{2-3}$ as numeric nutrient criteria. Benthic (i.e., stream-bottom) algae level criteria for these ecoregions should be set at $\leq$ 150 mg Chl $a$/m$^2$ ($\leq$ 36 g AFDW/m$^2$). The nutrient criteria should be established as the 80$^{th}$ percentile of the applicable level IV ecoregion reference distribution ONLY IF the screening conditions listed in Section 7.4 are met. If not, the nutrient and algae criteria for these level IV ecoregions should continue to be the same as for their parent level III ecoregion.

- DEQ should continue sampling reference streams in order to better characterize regional reference conditions and, in turn, refine the nutrient criteria. In particular, more nutrient samples are needed in the transitional mountain-to-prairie ecoregions. It is also recommended that another stressor-response study be carried out in the prairie regions of Montana, since there is only one completed at this time.
Appendix A

Development of Numeric Nutrient Criteria for Montana Prairie Streams (Northwestern Glaciated Plains Ecoregion)

Abstract

Twenty-four sites on twenty-two prairie streams of the Northwestern Glaciated Plains ecoregion in Montana were sampled from 2001-2004 in order to help establish regional nutrient criteria. Two approaches to deriving nutrient criteria were integrated: the stressor-response approach (nutrient as stressor, impact to beneficial water use as response), and the reference approach. Stream sites manifesting a range of conditions from excellent to poor were chosen to establish a human-disturbance gradient. Short stream reaches were sampled multiple times from late May to late September for a suite of parameters including riparian habitat condition, channel morphology, nutrient concentrations, aquatic plant biomass, and diatom-algae communities. Eight sites were determined to be in a reference condition. Data were used to determine the key parameters affecting the streams’ aquatic plant communities (i.e., we identified key ecological drivers). Several diatom metrics were calculated and subjected to five screening conditions before they would be considered for use in nutrient criteria development. These were (1) the metric had to be linkable to a numeric Montana water quality standard, (2) there was a significant correlation between the metric and a nutrient concentration, (3) the metric changed (increased, decreased) in the direction expected a priori, (4) reference site data points in the scatterplot between the metric and a nutrient were located in the region of the plot where they were expected, and (5) the metric was fairly insensitive to other important variables measured in the streams. Vulnerability to scouring flows and TSS concentrations were found to be major driving variables in the streams and influenced the aquatic plant communities observed. One diatom metric, the Oxygen Tolerance Index (OTI), met all 5 screening conditions and was used to derive a nutrient criterion. Three key characteristics of the metric were (1) it was insensitive to changes in TSS and measures of stream-scour potential, those variables shown to have great influence in the streams, (2) it did not respond to changes in total benthic plant biomass (benthic algae & macrophytes), or streambed macrophyte cover, and (3) it was responsive to nutrient concentration gradients after the elimination of the confounding influence of organic pollution present in some streams. The OTI metric was used to infer the streams’ nighttime DO concentrations, which were in turn compared to state minimum DO standards. The OTI was positively correlated to TN concentrations and a significant changepoint (p = 0.026) occurred at 1.12 mg TN/L. Streams having TN concentrations below the changepoint were, on average, in compliance with the applicable DO standard (5 mg/L), while streams with concentrations greater than 1.12 mg TN/L were not. All of the reference sites in the diatom OTI vs. TN concentration scatterplot were located to the left of the changepoint where one would expect them to be.

November, 2008  52
Appendix A

Section 1.0
Introduction

The over enrichment of waterbodies by nutrients (usually nitrogen [N] and phosphorus [P]) can increase nuisance algal growth, alter aquatic communities, and result in undesirable water-quality changes that impair beneficial water uses such as fisheries & aquatic life, irrigation, and water-supply\textsuperscript{23,32,33,55,102}. The U.S. Environmental Protection Agency (EPA) has published a series of documents containing ecoregion-specific nutrient criteria recommendations that are intended to control nutrient over-enrichment problems in streams and lakes (e.g., recommendations for streams in nutrient ecoregion IV\textsuperscript{103}) and, in turn, protect beneficial uses. The EPA has indicated that these criteria are preliminary, however, and work remains to develop more region-specific, scientifically-based nutrient criteria\textsuperscript{52,58}. Two approaches recommended by EPA to develop numeric nutrient criteria to protect aquatic systems are the reference approach and the stressor-response approach\textsuperscript{58}.

The reference approach relies on the identification of relatively undisturbed examples of waterbodies (i.e., reference sites\textsuperscript{86,104}), and proper geospatial classification of both the reference and non-reference waterbodies to assure “apple to apple” comparisons\textsuperscript{105,106}. This knowledge establishes a baseline against which changes in stream characteristics can be compared\textsuperscript{107,108}. More specific definitions such as ‘Minimally Disturbed Condition’, ‘Least Disturbed Condition’, etc. have been proposed to better classify waterbodies categorized as reference\textsuperscript{86}.

The stressor-response approach comprises two broad categories; studies carried out in laboratories, and those undertaken in the field. The most controlled stressor-response techniques are laboratory concentration-response studies between aquatic organisms and aqueous concentrations of a toxin\textsuperscript{109,110}, and EPA has well-developed protocols for these methods\textsuperscript{111}. But the focus of this discussion is on the other category of stressor-response studies (those using field data), which seeks to define quantitative relationships between an ecological response parameter (e.g., a stream fish population) and a gradient of an environmental condition\textsuperscript{112}. In the context of the present study, nutrient concentrations are the environmental condition of interest. This field-based approach has elsewhere been referred to as a mensurative experiment\textsuperscript{81}. Algae are commonly used for developing these field-based, stressor-response relationships. For example, regression equations are developed between benthic algal biomass and stream nutrient concentrations\textsuperscript{55,57,113}. Diatom algae in particular have been used for a long time as indicators of nutrient/organic enrichment in European rivers\textsuperscript{97,114-116}, and are also widely used in North America\textsuperscript{117-121}. Work in Europe shows that, along a human-caused gradient of stream condition, diatoms have high discriminatory power along the gradient and are more strongly correlated to eutrophication than are macrophytes, macroinvertebrates, or fish\textsuperscript{122,123}. Others have developed models relating diatom algae to particular nutrients such as total N and phosphate\textsuperscript{120,124,125}.

Reference and stressor-response approaches are discussed as separate (though complementary) techniques for developing nutrient criteria in EPA guidance\textsuperscript{58}. However, to establish a “gradient of environmental condition”\textsuperscript{112}, it is usually necessary to locate stream sites from the minimally to the highly impacted ends of an environmental spectrum, and those at the minimally-impacted end often meet at least some general definition of reference; in effect, this integrates the reference and stressor-response approaches. Several studies specifically involving diatom algae...
integrate reference and stressor-response techniques because “reference” or “minimally impacted” sites are used to establish the high-quality end of the gradient\textsuperscript{118,119,123,126}. But it is often the goal of these studies to develop general indices of stream health or biological integrity, and to develop nutrient criteria there still remains the dual tasks of (1) linking the stream-health indices or metrics to specific nutrients and (2) deciding where along the ecological condition gradient an impact to a beneficial water use has occurred. The first task is certainly feasible\textsuperscript{28,78,124,125}, but the latter can be difficult. If one were to use only the “pristine” (if available) end of the spectrum to establish nutrient criteria, criteria setting would be fairly easy. But setting criteria at pristine raises issues of cost, plausibility and public acceptance, and once one attempts to set criteria at something other than pristine difficult questions arise, such as “how much benthic algae is too much?” and “how much change in the macroinvertebrate community is acceptable?” Such questions are not easily answered. Answering these questions requires value judgments as well as scientific understanding, and gets directly at what has been termed “valued ecological attributes”\textsuperscript{3}, defined as ecosystem characteristics that directly or indirectly contribute to human welfare.

Our purpose in preparing this appendix is to describe how we used water quality tolerances of diatom assemblages to help derive nutrient criteria for a group of prairie streams in Montana. To do this, we first characterized physical, chemical, and habitat parameters of the streams in order to better understand which parameters had the most influence on aquatic plant biomass and diatom assemblages (i.e., what was the basic ecology of the streams). We were then able to establish linkages between a diatom assemblage water-quality index and stream nutrient concentrations. We used information provided by the diatom assemblage to infer stream dissolved oxygen (DO) concentrations, which allowed us to relate the stream nutrient concentrations to established state DO standards. Minimum stream DO requirements in state law (intended to protect aquatic life), reference sites, and sharp changes in the diatom index vs. nutrient concentrations were all used together to identify a threshold for harm to uses (i.e., fish and aquatic life). Thus, the stressor-response and reference approaches were integrated, which optimized the results and provided more confident conclusions.
Section 2.0
Methods

2.1 Description of Study Area

The study was carried out in prairie streams of Montana’s part of the Northwestern Glaciated Plains ecoregion\textsuperscript{63,64} (Figure 2.1), from 2001 to 2004. Prior to European settlement, the region was a semi-arid mixed prairie\textsuperscript{127}, and is now used mainly for grazing and growing cereal grain crops. In 2001 parts of the region were experiencing extreme summer drought based on the Palmer drought severity index\textsuperscript{128} housed at the NCDC (National Climate Data Center, historic monthly Palmer drought severity index records; available at \url{http://www.ncdc.noaa.gov}). Drought had been ongoing in Montana since 1998. During the 2002-2004 period, the summertime Palmer index returned to normal or near-normal ratings as precipitation improved.

![Figure 2.1. Sites Sampled in the Study. The Northwestern Glaciated Plains ecoregion is shaded.](image)

2.2 Study Site Selection

Sites were selected along an estimated “human disturbance” gradient, from best condition (reference) to most impacted. Sampling sites were selected on the basis of land cover information, field observations during site visits, and best professional judgment. In 2001, prior
Appendix A

to fieldwork, candidate sites were selected using a Geographic Information System (GIS) assisted approach. Fifth-field hydrologic unit code basins (HUCs\textsuperscript{129}) were overlaid in GIS with Multi-Resolution Land Characteristics (MRLC) land-use data and Strahler stream order\textsuperscript{4}. The proportion of agricultural land, forested land, urban land, and natural land cover were calculated for the 3\textsuperscript{rd} and 4\textsuperscript{th} order streams. A simple human-impact rating was assigned to each land type; for example, agricultural lands were rated as “high impact”, natural land cover was rated as “low impact”. Candidate stream sites where ranked (best to worst) based on the ratings and the proportion of each land type in their basins. Field reconnaissance in spring 2001 was undertaken to evaluate the overall condition of each candidate site. Best professional judgment was used to make qualitative evaluations of the condition of the riparian plant community, the degree of grazing impact, streambed stability, bank erosion, and human-caused dewatering. Sites to be sampled were chosen only after visiting many candidate sites in the ecoregion. In total, 24 sites on 22 streams were selected.

2.3 Final Site Rankings and Identification of Reference Stream Sites

After the 24 sampling sites were selected, detailed on-site data were collected which better enabled ranking of the sites’ conditions. (A description of on-site data collection procedures is presented later in Methods.) Each site was scored using three different stream condition assessment procedures: MT DEQ’s assessment form (1995-version), EPA’s rapid habitat assessment form\textsuperscript{130}, and EPA’s Environmental Monitoring and Assessment Program riparian disturbance metric (W1\_HALL\textsuperscript{131}). These scores were normalized to a 0-100 scale and then averaged (Table 2.1). Each procedure employs semi-quantitative visual estimates of various human impacts at each stream site, estimates that are ultimately converted to a quantitative score.

MT DEQ has screening criteria and procedures for identifying reference sites\textsuperscript{42}, and these were applied to study sites for the purpose of identifying those in a reference condition. About half of the 24 sites scored low enough on the composite stream-condition assessments discussed above that it was unlikely that they were in a reference condition. Therefore, the reference evaluation process was applied only to streams in the upper half of the score-ranked sites (Table 2.1). Eight sites passed all screening criteria, and are considered reference sites\textsuperscript{42} (Table 2.1). The eight reference sites are not pristine, but better fit the Least Disturbed Condition definition\textsuperscript{86}.
Reach Layout, Stream Habitat Assessment, and Sampling Frequency

Each site was laid out as a short reach determined as 40 times the mean wetted width taken at the initial visit, or a minimum of 150 m of stream length, and assessed using EPA’s Environmental Monitoring and Assessment Program (EMAP) physical habitat characterization protocols. Other stream condition and human impacts were assessed as well (see Section 2.0 of the Northwestern Glaciated Plains technical report by Supplee). Each stream site was sampled from 2 to 4 times per year during a restricted time frame between late May and late September. This period will be referred to hereafter as the “spring-summer period”. Repeat visits to each site during the spring-summer period were spaced so that they occurred approximately thirty days apart. The study was carried out over a four-year period, not all sites were sampled in each year, and some sites were sampled multiple years (Table 2.2). Reference sites were sampled most frequently, as we wanted to characterize them the most thoroughly given our available resources. At sites sampled over multiple years, physical habitat characterizations were done only in the first year since stream channel/riparian conditions did not substantially change from 2001 to 2004.
Tabel 2.2 Inventory of Repeat Measures for Parameters Measured at Each Site, 2001-2004.

<table>
<thead>
<tr>
<th>Stream Site Name</th>
<th>Year(s) Sampled</th>
<th>Reference Site (y/n)</th>
<th>Nutrient Samples</th>
<th>Common ions, TSS, pH, etc.</th>
<th>Flow</th>
<th>Total Benthic Plant Biomass*</th>
<th>Diatom Assemblage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock Cr (BLM)</td>
<td>2004</td>
<td>Yes</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td>Clear Creek</td>
<td>2001, 2003</td>
<td>Yes</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>21</td>
<td>6</td>
</tr>
<tr>
<td>Bitter Cr</td>
<td>2004</td>
<td>Yes</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>31</td>
<td>6</td>
</tr>
<tr>
<td>Rock Creek (Site 1)</td>
<td>2003</td>
<td>Yes</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Willow Cr (North)</td>
<td>2001, 2002</td>
<td>Yes</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>36</td>
<td>5</td>
</tr>
<tr>
<td>Wolf Creek at Wolf Point</td>
<td>2002, 2003, 2004</td>
<td>Yes</td>
<td>9</td>
<td>9</td>
<td>4</td>
<td>49</td>
<td>15</td>
</tr>
<tr>
<td>Rock Creek (Site 2)</td>
<td>2001, 2002, 2003, 2004</td>
<td>Yes</td>
<td>12</td>
<td>12</td>
<td>9</td>
<td>66</td>
<td>17</td>
</tr>
<tr>
<td>Sheep Cr</td>
<td>2002</td>
<td>No</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>12</td>
<td>4</td>
</tr>
<tr>
<td>Horse Tied Cr</td>
<td>2002</td>
<td>No</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>Porcupine Cr</td>
<td>2001, 2002, 2004</td>
<td>No</td>
<td>11</td>
<td>10</td>
<td>7</td>
<td>70</td>
<td>12</td>
</tr>
<tr>
<td>Shotgun Cr</td>
<td>2002</td>
<td>No</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>Middle Fork Poplar River</td>
<td>2002, 2004</td>
<td>No</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>54</td>
<td>10</td>
</tr>
<tr>
<td>Wolf Creek nr Medicine Lake</td>
<td>2002</td>
<td>No</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>22</td>
<td>3</td>
</tr>
<tr>
<td>Big Sandy Cr</td>
<td>2001, 2004</td>
<td>No</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>55</td>
<td>8</td>
</tr>
<tr>
<td>Larb Cr</td>
<td>2001</td>
<td>No</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>Butte Cr</td>
<td>2002, 2004</td>
<td>No</td>
<td>7</td>
<td>8</td>
<td>5</td>
<td>53</td>
<td>13</td>
</tr>
<tr>
<td>Beaver Cr</td>
<td>2001</td>
<td>No</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Frenchman Cr</td>
<td>2001</td>
<td>No</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Battle Cr</td>
<td>2001</td>
<td>No</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>Redwater River</td>
<td>2002</td>
<td>No</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>21</td>
<td>4</td>
</tr>
<tr>
<td>Smoke Cr</td>
<td>2002</td>
<td>No</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>21</td>
<td>5</td>
</tr>
<tr>
<td>Little Muddy Cr</td>
<td>2002</td>
<td>No</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>Willow Cr (South)</td>
<td>2001, 2004</td>
<td>No</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>53</td>
<td>6</td>
</tr>
</tbody>
</table>

* Each sample collected at each transect is separately inventoried here. See text for details on reach layout and sample collection.

### 2.5 Water Quality Measurements

See Section 2.3.2 and 2.3.3 in Suplee\textsuperscript{41} for details on water quality sampling of nutrients, common ions, etc. Flow was measured during site visits using the velocity-area method\textsuperscript{132}. Real-time water quality measurements were taken during each field visit including pH, specific conductance (µS/cm @ 25°C), temperature, and dissolved oxygen (DO). To provide a cross-check of temperature measurements taken during field visits, continuous temperature measurements were collected using Optic Stow Away\textsuperscript{®} loggers placed in three sites (Rock Cr [Site 2], Porcupine Cr and Wolf Cr at Wolf Point) during summer 2004. Mean summer-long temperatures derived from the loggers placed in these three streams were 21, 19 and 21°C, respectively. Stream water DO at saturation was calculated using the mean elevation of the sites (713.5 m; 1 standard error = 25 m) and the mean spring-summer water temperature (from field measurements) for all sites (22°C).

### 2.6 Algal and Macrophyte Biomass Sampling (Quantitative)

See Section 2.3.4 of Suplee\textsuperscript{41} for details on macrophyte, filamentous algae, and phytoplankton sampling and laboratory analysis. A change that occurred for 2004 was the separate collection of macrophytes and filamentous algae. Whereas in the first years of the project all aquatic plant material in a given hoop sample (macrophytes, filamentous algae) was analyzed together for reporting of Chl \textit{a} and AFDW, starting in 2004, individual hoop samples were processed in the
field such that the macrophyte component and the filamentous algae component were physically separated so they could later be quantified individually. In this appendix the term “total benthic plant biomass” refers to the aggregate biomass, in mg Chl/ m^2, of floating filamentous algae, attached benthic algae, and macrophytes. It does not include phytoplankton biomass, which were separately measured and are expressed as µg Chl /L.

2.7 Determination of Algal Taxonomical Composition

Composite samples from each site during each visit were collected and preserved (Lugol’s solution) for identification of soft-bodied algae and diatoms. In 2001 and 2002, composite samples of three algae habitat types were collected: “plant” type, “rock” type and “sediment” type. After collection of the quantitative Chl a samples at each transect, a qualitative sample of the same representative material was also collected. For the “sediment” type samples, only material from the very surface of the stream bottom was collected. For each reach, all material from a common sample type (e.g., all “plant” type subsamples) was composited, resulting in up to three composite habitat-type samples per stream site, per visit. In 2003 and 2004, multi-habitat composite samples were collected from each site and comprised all three habitat types described above. Samples were submitted to Hanaea in Helena, MT for identification of soft-bodied algae and diatoms. Relative abundance and ordinal rank by biovolume of diatoms and genera of soft (non-diatom) algae were determined per methods by Bahls. Algae of the division Cyanophyta (cyanobacteria) were also identified. Soft algae were identified using standard taxonomic texts. A subsample of each sample was then cleaned of organic matter using potassium dichromate, sulfuric acid, and hydrogen peroxide. Permanent diatom slides were prepared using Naphrax following standard methods. Approximately 300 diatom cells (600 valves) were randomly counted and identified to species using established taxonomic references. Diatom naming conventions followed those adopted by the Academy of Natural Sciences for USGS NAWQA samples.

Published diatom indexes (i.e., metrics) were reviewed to see if they could be related to Montana’s water quality standards. By metric we mean a group of diatom attributes used together to assess a water quality condition, for example to tell us about pH. Metrics considered were (1) a metric for pH and (2) an oxygen tolerance index (OTI). The OTI classes range from 1 to 5 and—counter intuitively—higher class values indicate decreasing DO concentrations. That is, the oxygen tolerance index is assessing tolerance to low DO conditions. For each sample, a weighted-average OTI score was calculated by multiplying each DO tolerance value or class (i.e. 1, 2, 3, 4, or 5) by the proportion of sample taxa in that class, summing these results, and then dividing by the proportion of OTI-classified taxa in the entire sample. If, for example, 25% of the taxa in a sample were not classified into one of the five OTI classes, the denominator would be 0.75.

We also wanted to make distinctions between DO saturation-deficit resulting from eutrophication vs. DO saturation-deficit attributable mainly to organic pollution (i.e., carbonaceous or nitrogenous BOD). For each sample, the proportion of organic pollution tolerant taxa in each sample was determined per Kelly and Whittone and Kelly. Sites with >
40% organic pollution tolerant taxa are likely to have significant organic pollution\textsuperscript{124}, and such sites were removed from the dataset in order to carry out analyses (e.g., correlation with nutrients) without the confounding influence of organic pollution. Table 2.3 provides the diatom taxa considered to be organic pollution tolerant.

Table 2.3. List of Organic Pollution Tolerant Diatoms\textsuperscript{*}, Per Kelly and Whiton\textsuperscript{124} and Kelly\textsuperscript{145}.

\begin{center}
\begin{tabular}{l}
\textit{Gomphonema parvulum} \\
\textit{Navicula gregaria} \\
\textit{Navicula incertata} \\
\textit{Navicula lanceolata} \\
\textit{Navicula minima} \\
\textit{Navicula pelliculosa} \\
\textit{Navicula saprophila} \\
\textit{Navicula subminuscula} \\
\textit{Navicula tenelloides} \\
\textit{Nitzschia spp.} \\
\textit{Sellaphora seminulum}
\end{tabular}
\end{center}

*List was crossed-check with L. Bahls for applicability to Montana taxa autecology.

Biological metrics are generally designed around population-level changes, and the dominance of a single or a few unclassified taxa in a sample tends to compromise the sample’s metric results\textsuperscript{120} (also W. Bollman, personal communication, April 2008; D. Charles, personal communication, April 2008). “Unclassified” taxa are those that do not play a role in the metric in question. Nine samples (of 152) were eliminated from the OTI metric dataset because a solid majority (> 55%) of each of the sample’s diatom counts was given as “unclassified”; that is, the majority of diatom species encountered did not fall into any of the five OTI tolerance classes. Taxa counts from each of these nine samples were reviewed, and in all nine samples diatom abundance was dominated by a few or even a single taxa. Further, the total number of taxa in most of these samples was greatly reduced relative to what was seen in samples from the rest of the study. The usability of the nine samples was likely compromised by a reduced overall population and/or the dominance of a few unclassified taxa, and therefore were not used in the analyses.

2.8 Data Analyses, Statistical Tools, and Data Compilation

To understand differences between reference and non-reference streams and to assess the influence of different stream parameters on aquatic plant measurements, we carried out statistical tests of difference and correlation. We examined parameters that have already been shown to affect aquatic plant growth in streams (for overview see Wetzel\textsuperscript{9} and Stevenson \textit{et al.}\textsuperscript{36}) and parameters that helped us evaluate the usefulness of the diatom metrics (more on this, next subsection). Parameters important to stream aquatic plants include light\textsuperscript{146-148}, nutrients\textsuperscript{56,62,83,148,149}, stream scour/TSS\textsuperscript{45,150-154}, substrate\textsuperscript{155,155}, and total dissolved salts\textsuperscript{156,157}.
Each study site was considered to be independent. Most sites were located on different streams, but those on the same stream (i.e., the three Rock Creek sites) were considered independent because (1) there was at least 1.6 km of stream distance between them and (2) at least one tributary confluence was found between the sites. Sampling frequency varied greatly in this study. Between 2001 and 2004 some sites were sampled only in one year, while others were sampled for two, three or even four years (Table 2.2). During any given spring-summer period, each site was repeatedly sampled from 2 to 4 times. To preclude potential temporal pseudoreplication issues, data collected at each site were reduced to means. For example, if a site was sampled 4 times during the spring-summer period but only in 2002, the 4 repeat measures of parameter \( X \) were used to calculate the mean; if another site was sampled 3 times in 2001 and 4 times in 2004, all 7 repeat measures of \( X \) were used to calculate the site mean. This resulted in an \( n \) of 24 (8 reference, 16 non-reference sites) for each parameter, one value per site, with some sites being better characterized over the study period than others.

Statistical analyses were made using Minitab® (Release 15). Statistical differences were considered significant when \( p \)-values were < 0.10. For all parameters, non-detects were replaced with values equal to 50% of the reported detection limit prior to use in statistical tests. One-sided tests were used in all situations where a priori relationships were expected. The Mann-Whitney test was used to determine significant differences between reference and non-reference sites. The Spearman Rank test was used to examine the strength (rho) and significance (p-value) of correlations between parameters. Spearman Rank p-values were not taken from Minitab® (which Minitab® states are inaccurate), but were instead calculated as shown on page 317 of Conover. In the correlations we did not apply a Bonferroni adjustment to the 0.1 significance threshold because (1) each smaller “study” (i.e., analysis) done in the context of the larger study was considered on its own merits, and (2) we did not want to further increase the chance of making type II errors (i.e., declaring truly significant relationships insignificant).

Changepoint analysis was used to identify the occurrence of thresholds in scatterplots between diatom metrics and nutrients. Language was written for Stata® (version 10) to carry out changepoint analysis via the deviance reduction approach. The analysis included an approximate Chi-square-test to determine the significance of the changepoint. The Chi-square test assumes that the deviance reduction divided by the scale parameter is approximately Chi-square distributed.

### 2.9 Identifying a Harm-to-Beneficial Use Threshold

Five conditions were used to assess whether or not a diatom metric would be useful for helping establishing nutrient criteria. These were (1) the metric had to be meaningfully linked to a numeric Montana water quality standard, (2) there was a significant correlation between the metric and a nutrient concentration, (3) the metric changed (increased, decreased) in the direction expected a priori, (4) reference site data points in the scatterplot between the metric and a nutrient needed to be located in the region of the plot where they were expected, and (5) the metric was fairly insensitive to other important variables measured in the streams.

Montana water quality standards include numeric criteria for DO and pH, and diatom metrics that could be linked to these standards were described earlier (see Section 2.7, this Appendix). The pH metric was not further considered, however, because the streams all had pH values $\geq 7.6$. 

November, 2008
Appendix A

(maximum 9.0), which only permitted the use of one or at most two (of six) pH tolerance-values in van Dam et al. In effect, the attribute could not be evaluated by condition No. 3 above.

For the diatom oxygen tolerance index (OTI), van Dam et al. provide diatom-assemblage tolerance values (classes) that correspond with DO saturation deficits (e.g., diatoms with a tolerance value of 3 tolerate DO between 75% and 50% of saturation, while tolerance-value 2 diatom taxa are found where DO is between 100% and 75% of saturation). A simple linear regression was made between van Dam et al.'s five tolerance values (x axis) and their associated % DO saturation requirements (y axis) (Figure 2.2). The linear regression could then be used to convert metric scores to inferred % DO saturation.

\[
y = -0.2275x + 1.2825 \\
R^2 = 0.98
\]

Figure 2.2. Relationship Between Diatom Metric Values and Inferred Percent DO Saturation of Stream Water (After van Dam et al.).

DO concentrations were then estimated for each sample by multiplying the diatom-inferred % saturation by the stream’s DO at saturation. Diatom-inferred DO concentration estimates were then compared to state DO standards. Montana DO standards vary by averaging period (e.g., 1 day, weekly, monthly, etc.). Given the typical generation time of diatoms, the DO estimates from the aggregate diatom-community samples are best associated with the weekly DO standards (also L. Bahls, personal communication, April 2006), as opposed to the monthly or daily standards. Therefore, we compared the diatom-inferred DO estimates to the state's 7 Day Mean Minimum standard. The 7 Day Mean Minimum is the average of each of the lowest daily DO values measured over a seven day period; these minima almost always occur at night.
Section 3.0
Results

3.1 Physical, Chemical, and Biological Characteristics of the Study Streams

Riparian habitats ranged from those dominated by sedges and rushes and having no trees and few shrubs, to deciduous forests of green ash, boxelder, and American elm (Table 2.1). The most common riparian habitats were the Shrub and Herbaceous Riparian complex and the Recent Riparian Complex\textsuperscript{163}; this was true for both reference and non-reference sites. Flow was highly variable (note minimum and maximums; Table 3.1), and most sites ceased flowing in summer and became intermittent (i.e., a series of disconnected pools). Half the reference sites could be considered perennial and half intermittent, while 31\% of non-reference sites were perennial and 69\% intermittent. Individual TSS samples varied over four orders of magnitude across all sites, while TSS site means ranged from 3 (West Fork Poplar River) to 513 mg/L (Willow Cr [South]). TSS correlated significantly with Rosgen entrenchment ratio ($\rho = -0.311$, $p = 0.077$). Higher Rosgen entrenchment ratios are found in streams having better-developed floodplains, inferring decreased potential for streambed scour\textsuperscript{164}. Thus, the streams with higher scouring potential had higher TSS concentrations. The streams generally had high ionic concentrations, as mean specific conductivity for most streams was above 1000 $\mu$S/cm during the spring-summer period. The dominant cation and anion were $\text{Na}^+$ and $\text{SO}_4^{2-}$. 

November, 2008  63
TN concentrations were significantly lower in reference sites than in non-reference sites (p-value = 0.03). Reference sites had a mean TN of 863 µg/L and 1 standard error of the mean equal to 58 µg/L (mean ± 1SE), while non-reference sites had a mean of 1241 µg/L (1 SE = 498 µg/L; Table 3.1). Among the 24 sites, mean NO₂⁺³ concentrations varied over three orders of magnitude (0.3 to 699.7µg/L) and were more variable in the non-reference sites than the reference sites (Table 3.1); however, no significant difference between the groups was detected (test assumption that reference NO₂⁺³ < non-reference was not fulfilled). TP concentrations in reference sites were not significantly lower than non-reference sites (p-value = 0.46). Similarly, mean SRP in the reference sites was not significantly lower (p = 0.24) than in the non-reference sites (Table 3.1).

Total benthic plant biomass in the reference sites averaged 58 mg Chl a/m² (1 SE = 20), and in the non-reference sites 104 mg Chl a/m² (1 SE = 22; Table 3.1). There was no significant difference in total benthic plant biomass between the two groups (p = 0.3, two-sided test). The % stream bottom macrophyte cover averaged 14% (1 SE = 9%) in the reference sites and 30% (1 SE = 7%) in the non-reference sites, and there was no significant difference between the two groups (p = 0.11, two-sided test). Finally, phytoplankton density averaged 9.9 µg/L (1 SE = 3 µg/L) in the reference sites and 39.9 µg/L (1 SE = 29 µg/L) in the non-reference sites, and there was no significant difference between the reference and non-reference sites (p = 0.91, two-sided test).
3.2 Response of Non-diatom Aquatic Plant Measurements to Environmental Parameters

Total benthic plant biomass was negatively correlated with TSS (rho = -0.58), and positively correlated with Rosgen entrenchment ratio (rho = 0.39; Table 3.2) — note that higher entrenchment ratios are associated with channels that are better able to access wide floodplains. Thus, more total benthic plant biomass was found in sites that had a reduced likelihood of scouring flows. Riparian canopy density, stream substrate size, and electrical conductivity were not significantly correlated with total benthic plant biomass. TP was negatively correlated with total benthic plant biomass (rho = -0.48), however TP was also significantly correlated (positively) with TSS (rho = 0.53), suggesting that TP may have been acting as a TSS surrogate in these streams. Total N was not significantly correlated with TSS (Table 3.2). The % streambed cover by macrophytes correlated negatively with TSS and TP (rho = -0.63 and -0.39, respectively; Table 3.2) but not with any other parameters. Surprisingly, phytoplankton density — quantified as µg Chl a/L — correlated positively with TSS (rho = 0.59). Phytoplankton correlated with several nutrients as well (TN, TP, and NO$_2$+3).
### 3.3 The Diatom OTI in Relation to Environmental Parameters

Using the complete dataset (24 sites), it was found that the diatom OTI did not correlate to Rosgen entrenchment ratio, TSS, total benthic plant biomass, or % macrophyte cover, but did

---

#### Table 3.2: Spearman Rank Correlations (rho) and p-values for Correlations Between Driving and Response Variables. Significant Relationships (p < 0.1) Shown in Bold.

<table>
<thead>
<tr>
<th>Response Variable</th>
<th>Driving Variable</th>
<th>Expected Relationship</th>
<th>Spearman's rho</th>
<th>n</th>
<th>Estimated p-value, from Table A1 of Canover (1999)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>Bank canopy density</td>
<td>negative</td>
<td>0.041</td>
<td>23</td>
<td>0.424</td>
</tr>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>Midchannel canopy density</td>
<td>negative</td>
<td>0.115</td>
<td>23</td>
<td>0.295</td>
</tr>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>Stream substrate D0</td>
<td>not sure</td>
<td>0.273</td>
<td>23</td>
<td>0.732</td>
</tr>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>Log10 mean TSS conc. (mg/L)</td>
<td>negative</td>
<td>0.577</td>
<td>23</td>
<td>0.004</td>
</tr>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>Rosgen entrenchment ratio</td>
<td>positive</td>
<td>0.386</td>
<td>22</td>
<td>0.038</td>
</tr>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>EC (us/cm)</td>
<td>negative</td>
<td>0.212</td>
<td>23</td>
<td>0.160</td>
</tr>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>TN conc. (mg/L)</td>
<td>positive</td>
<td>0.266</td>
<td>23</td>
<td>0.106</td>
</tr>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>NO₃ conc. (mg/L)</td>
<td>positive</td>
<td>0.202</td>
<td>23</td>
<td>0.172</td>
</tr>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>TP conc. (mg/L)</td>
<td>not sure</td>
<td>0.475</td>
<td>23</td>
<td>0.026</td>
</tr>
<tr>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>SRP conc. (mg/L)</td>
<td>positive</td>
<td>0.003</td>
<td>18</td>
<td>0.449</td>
</tr>
<tr>
<td>Streamed macrophyte cover (%)</td>
<td>Midchannel canopy density</td>
<td>negative</td>
<td>0.372</td>
<td>24</td>
<td>0.100</td>
</tr>
<tr>
<td>Streamed macrophyte cover (%)</td>
<td>Stream substrate D0</td>
<td>not sure</td>
<td>0.234</td>
<td>24</td>
<td>0.290</td>
</tr>
<tr>
<td>Streamed macrophyte cover (%)</td>
<td>Log10 mean TSS conc. (mg/L)</td>
<td>negative</td>
<td>0.373</td>
<td>24</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Streamed macrophyte cover (%)</td>
<td>Rosgen entrenchment ratio</td>
<td>positive</td>
<td>0.14</td>
<td>22</td>
<td>0.261</td>
</tr>
<tr>
<td>Streamed macrophyte cover (%)</td>
<td>EC (us/cm)</td>
<td>negative</td>
<td>0.352</td>
<td>24</td>
<td>0.056</td>
</tr>
<tr>
<td>Streamed macrophyte cover (%)</td>
<td>TN conc. (mg/L)</td>
<td>not sure</td>
<td>0.12</td>
<td>24</td>
<td>0.338</td>
</tr>
<tr>
<td>Streamed macrophyte cover (%)</td>
<td>NO₃ conc. (mg/L)</td>
<td>positive</td>
<td>0.379</td>
<td>24</td>
<td>0.070</td>
</tr>
<tr>
<td>Streamed macrophyte cover (%)</td>
<td>TP conc. (mg/L)</td>
<td>not sure</td>
<td>0.389</td>
<td>24</td>
<td>0.062</td>
</tr>
<tr>
<td>Streamed macrophyte cover (%)</td>
<td>SRP conc. (mg/L)</td>
<td>positive</td>
<td>0.022</td>
<td>18</td>
<td>0.464</td>
</tr>
<tr>
<td>Phytoplankton concentration (mg/L)</td>
<td>Midchannel canopy density</td>
<td>negative</td>
<td>0.361</td>
<td>22</td>
<td>0.230</td>
</tr>
<tr>
<td>Phytoplankton concentration (mg/L)</td>
<td>Stream substrate D0</td>
<td>not sure</td>
<td>0.335</td>
<td>22</td>
<td>0.124</td>
</tr>
<tr>
<td>Phytoplankton concentration (mg/L)</td>
<td>Log10 mean TSS conc. (mg/L)</td>
<td>positive</td>
<td>0.368</td>
<td>22</td>
<td>0.005</td>
</tr>
<tr>
<td>Phytoplankton concentration (mg/L)</td>
<td>Rosgen entrenchment ratio</td>
<td>negative</td>
<td>0.179</td>
<td>21</td>
<td>0.212</td>
</tr>
<tr>
<td>Phytoplankton concentration (mg/L)</td>
<td>EC (us/cm)</td>
<td>negative</td>
<td>0.031</td>
<td>22</td>
<td>0.443</td>
</tr>
<tr>
<td>Phytoplankton concentration (mg/L)</td>
<td>TN conc. (mg/L)</td>
<td>positive</td>
<td>0.252</td>
<td>22</td>
<td>0.053</td>
</tr>
<tr>
<td>Phytoplankton concentration (mg/L)</td>
<td>NO₃ conc. (mg/L)</td>
<td>positive</td>
<td>0.346</td>
<td>22</td>
<td>0.056</td>
</tr>
<tr>
<td>Phytoplankton concentration (mg/L)</td>
<td>TP conc. (mg/L)</td>
<td>positive</td>
<td>0.436</td>
<td>22</td>
<td>0.023</td>
</tr>
<tr>
<td>Phytoplankton concentration (mg/L)</td>
<td>SRP conc. (mg/L)</td>
<td>positive</td>
<td>0.121</td>
<td>18</td>
<td>0.309</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>Total benthic plant biomass (mg Chl a/m²)</td>
<td>positive</td>
<td>0.021</td>
<td>23</td>
<td>0.461</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>Phytoplankton concentration (mg/L)</td>
<td>positive</td>
<td>0.369</td>
<td>22</td>
<td>0.046</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>Streamed macrophyte cover (%)</td>
<td>positive</td>
<td>0.103</td>
<td>24</td>
<td>0.311</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>Log10 mean TSS conc. (mg/L)</td>
<td>not sure</td>
<td>0.188</td>
<td>24</td>
<td>0.183</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>Rosgen entrenchment ratio</td>
<td>not sure</td>
<td>-0.352</td>
<td>22</td>
<td>0.108</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>TN conc. (mg/L)</td>
<td>not sure</td>
<td>0.422</td>
<td>24</td>
<td>0.021</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>SRP conc. (mg/L)</td>
<td>positive</td>
<td>0.431</td>
<td>24</td>
<td>0.019</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>NO₃ conc. (mg/L)</td>
<td>positive</td>
<td>0.13</td>
<td>24</td>
<td>0.266</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>TP conc. (mg/L)</td>
<td>positive</td>
<td>0.446</td>
<td>24</td>
<td>0.016</td>
</tr>
<tr>
<td>Diatom OTI</td>
<td>Phosphorus concentration (mg/L)</td>
<td>positive</td>
<td>0.41</td>
<td>18</td>
<td>0.046</td>
</tr>
</tbody>
</table>

---

*Significant relationships that respond logic to the expected relationship are shown in bold.
1Combined biomass of benthic algae, macrophytes, and floating filamentous algae.
positively correlate with phytoplankton density (rho = 0.369). The diatom OTI correlated positively with several nutrients, including TN (Table 3.2). The scatterplot between the diatom OTI and TN is shown in Figure 3.1. Note the position of the reference sites on the left side of the scatterplot.

In the analyses up to this point, the diatom OTI had passed all 5 screening conditions. The OTI could (1) be linked to Montana water quality standards (DO), (2) it correlated significantly to several nutrients and (3) in the manner expected (Table 3.2), (4) the reference sites are generally where they would be expected to be (Figure 3.1), and (5) it was insensitive to other important non-nutrient factors in the streams (paragraph above). To further assess condition 5 (i.e., the metric is fairly insensitive to other environmental variables measured in the streams), only sites having <40% organic pollution tolerant taxa (n = 16 sites) were carried to the next step. The scatterplot for this diatom OTI vs. mean TN relationship is shown in Figure 3.2. A significant changepoint (p = 0.026) was detected at 1.12 mg TN/L where the mean diatom OTI increased from 2.5 (group of sites to the left of 1.12 mg TN/L) to 2.9 (sites to the right of 1.12 mg TN/L) (Figure 3.2). The 90% confidence interval for this changepoint is 0.78-1.48 mg TN/L. (Changepoint analysis was also run on the full [n=24] dataset. A significant changepoint [p <0.01] was also detected at 1.12 mg TN/L, with a 90% confidence interval from 0.84 to 1.3 mg TN/L.) Given that mean spring-summer temperature in the streams during the study was 22°C and the average site elevation was 713.5 m with little variance (1 SE = 25 m), stream DO at saturation was estimated to be 8 mg/L. Electrical conductivity (i.e., dissolved salts) was not high enough in the streams to affect DO saturation meaningfully.

Recall that diatom-inferred DO saturation equals: -0.2275·(OTI score) + 1.2825 (Figure 2.2). Hence, the group of streams with TN ≤ 1.12 mg TN/L (and a mean OTI score of 2.5) had a mean inferred DO of about 5.7 mg/L (-0.2275·2.5 + 1.2825 = 71%; 8 mg/L·71% = 5.7 mg DO/L). The group with TN > 1.12 mg TN/L had a mean inferred DO slightly less than 5.0 mg/L (-0.2275·2.9 + 1.2825 = 62%; 8 mg/L·62% = < 5.0 mg DO/L). None of the reference sites had a mean TN concentration greater than 1.12 mg TN/L, nor did any of them have an inferred DO concentration < 5 mg/L (Figures 3.1, 3.2).
Figure 3.1. Scatterplot of Diatom OTI vs. Total N Concentrations, All Sites. Diatom-inferred dissolved oxygen (DO) was calculated based on a DO at saturation of 8 mg/L.
Appendix A

Figure 3.2. Scatterplot of Diatom OTI vs. Total N Concentrations, Only Sites Where Organic Pollution Tolerant Taxa Were < 40% of the Population. Diatom inferred dissolved oxygen (DO) was calculated based on a DO at saturation of 8 mg/L.

Hollow: Reference  
Solid: Non-Reference  
Triangle: Macrophytes cover >10% of reach.
Appendix A

Section 4.0
Discussion

4.1 Important Driving Variables Affecting the Prairie Streams

Stressor-response relationships between nutrients and aquatic plants (including diatoms) are not easily studied in streams where multiple environmental variables are operating simultaneously. So, to the extent practical, we quantified stream nutrient concentrations as well as other, non-nutrient parameters and correlated these to measures of the aquatic plant community to better understand the trend and magnitude of their effects (Table 3.2). A complex but identifiable pattern begins to emerge when these data are examined together. Streams with the highest potential for bottom scour during high flows (i.e., those with low Rosgen entrenchment ratios) had higher TSS as well as lower total benthic plant biomass (Table 3.2). Fine sediments are common in prairie streams and are easily entrained during high flows, especially in the more entrenched channels. Thus, it would be expected that TSS and entrenchment relate to one another. Phytoplankton were positively correlated with TSS, which may at first glance be counter-intuitive as one might expect that high TSS would reduce light and, therefore, the phytoplankton standing crop. But light (measured as riparian canopy density) was apparently not a strong regulator of phytoplankton biomass, nor of benthic plant biomass or macrophyte cover (Table 3.2). Phytoplankton responded positively to TSS increases probably because elevated TSS was more common in scour-prone streams and scoured streams had fewer of the phytoplankton’s resource competitors (i.e., benthic plants).

In these prairie streams high-flow scour acted as a key driving variable and, depending on each stream’s propensity to become scoured, tended to result in two very different general endpoints (with intermediate forms also present); (1) relatively un-scoured streams where plant biomass was dominated by benthic forms, and (2) more heavily scoured streams where phytoplankton made up a major proportion of the algal biomass (see Section 6.0 and 8.0, and especially Figure 8.7, in Suplee). High flows are known to be an important regulator of benthic aquatic plant communities. And artificial stream studies show that an elevation in water velocity beyond that to which benthic algae are adapted leads to reduced biomass, whereas TSS additions in the same artificial streams do not have a marked effect on algal biomass. Flashy high flows are common in the region due to sporadic, heavy summer thunderstorms, and so the influence of scouring flow was likely to be manifested at most sites some time during the summer.

TP was negatively correlated with total benthic plant biomass and macrophytes and, following the same pattern as TSS, was positively correlated with phytoplankton. This most surely resulted because TP was acting as a TSS surrogate. The TP-TSS correlation is not unusual, as 90% of TP carried by rivers to the sea is associated with suspended solids. Due to the tight coupling of TSS and TP (Table 3.2) the role of TP as a plant nutrient vs. TP as a TSS surrogate is difficult to tease apart; therefore, TP will not be further addressed in this appendix as a plant nutrient.
4.2 Effects of Non-Nutrient Environmental Factors on the Diatom OTI

We saw a fairly clear pattern between the diatom OTI and TN concentrations (Table 3.2; Figures 3.1, 3.2). Because we would like to use the diatom OTI metric to infer DO concentrations and derive regional nutrient criteria, it is important to evaluate non-nutrient factors that might influence the metric given the complex nature of these streams. Macrophytes, for one, can influence diel DO concentrations but may not necessarily correlate to instream nutrients since they can get nutrients from the sediments via their roots. But neither macrophyte coverage nor total benthic plant biomass correlated with nutrients or the OTI metric in our study (Table 3.2), and note in Figure 3.1 that streams with substantial macrophyte cover (> 10%) show a wide range of diatom OTI values (i.e., macrophyte cover is not driving the diatom OTI values). Macrophytes are present in many of these streams, often in large quantities, but their role relative to the diatom OTI is more that of a random variable. The diatom OTI is not responding to TSS either, which is important since TSS has been shown to be a key driving variable in these streams. However, the OTI was responsive to phytoplankton, which were themselves closely linked to instream nutrients (TN and NO₂⁻³)(Table 3.2). Phytoplankton, every bit as much as benthic plants, can strongly affect diel DO concentrations oscillations. Lastly, the diatom OTI vs. TN relationship remains essentially unchanged after the sites likely contaminated with organic pollution are excluded (Figures 3.1, 3.2). Organic pollution and nutrient pollution often go hand in hand, but regardless, Figure 3.2 can be viewed as the best representation of the diatom OTI’s ability to respond to purely nutrient-caused DO variations in these streams. Overall, from the preceding discussion, it can be concluded that the diatom OTI is relatively insensitive to important non-nutrient environmental parameters in these streams (i.e., it meets condition No. 5; it also meets conditions 1 through 4), and is therefore useful for developing nutrient criteria.

4.3 The OTI relationship to Nutrient Concentrations, and the Dissolved Oxygen Estimates Derived from the Diatom OTI

Diatoms are known to be good indicators of stream TN concentrations and eutrophication in general. Other researchers find that individual tolerance classes of the diatom OTI metric correlate well to nutrient gradients in streams throughout North America, and in this particular region. Porter et al. report that among 35 different diatom attributes, including the 5 OTI classes of van Dam et al., class 4 of the OTI was one of the very best indicators of nutrient enrichment based on a U.S. national dataset from the USGS.

Important for the present work is an understanding of the nature of the DO estimates provided by the diatom OTI. The available data indicate that the DO values provided by the diatom OTI reflect nighttime minima. Six case studies (five in Montana, one from the literature) were available in which diatom samples were collected during the summer and instrument-measured 24 hr DO data were also available. Three case studies were from 2007 at two sites along the Yellowstone River (7th order large river), two were from Cottonwood Creekf (a 2-3rd order

---

f At Cottonwood Creek, DO data were measured 24 hr/day across two days at two sites in July 2003, while the diatom data were collected at the same two sites but in August 2001. Unfortunately, simultaneous diatom and DO instrument data were not collected. Multiple-year data from the sites suggest that stream conditions were generally comparable during the two time periods, therefore the 2003 DO data were compared to the 2001 diatom data.
wadeable stream), and the last was from The Grand River near Toronto, Canada; the latter is published in Rott et al.\textsuperscript{117}. Rott et al. use the saprobic index\textsuperscript{115} instead of the OTI. The saprobic index and the OTI provide very similar information about stream DO. Each index has five tolerance values, with analogous tolerance values in each index indicate essentially identical DO conditions\textsuperscript{115}. In our database, 69\% of the diatom species in analogous tolerance values of each index are identical. During summer, downstream sites in the Grand River have elevated N and P concentrations, instrument-measured DO deficits of 51\% of saturation, and instrument-measured nighttime DO concentrations < 4 mg/L. The elevation and summertime temperatures of the Grand River’s downstream sites equate to a DO concentration at saturation of around 8 mg/L. Most diatom samples from the Grand River’s downstream sites trend towards saprobic index class III, corresponding to 50\% DO saturation-deficit. Thus, diatom-inferred DO concentrations on the lower Grand River would equal about 4 mg/L, which corresponds well with the instrument-measured nighttime data.

We compared nighttime DO instrument data from each of the six case studies mentioned above to their corresponding diatom-inferred DO, in order to compute the diatom OTI's overall bias and relative error. The analysis showed that the diatom OTI has a slight low bias (it provides values that are 0.18 mg DO/L lower than "true" instrument-measured nighttime DO minima), with a standard error of the mean (SEM) around the DO estimate of ± 0.58 mg/L. DEQ is continuing to collect diatom and instrument DO data to further refine these measures of the OTI's bias and error.

4.4 Identifying the Threshold of Harm to the Beneficial Uses

Weekly DO standards vary in the study region depending upon classifications in state law. Most streams in the ecoregion are classified B-3 or C-3, meaning a DO standard of 4.0 mg/L. But there are also many B-1 and B-2 streams, which have a DO standard of 5.0 mg/L\textsuperscript{8}. In the diatom OTI vs. TN relationship (Figure 3.2), a clearly visible (and significant; p-value = 0.026) changepoint occurred at 1.12 mg TN/L. Streams with TN ≤ 1.12 mg TN/L had a mean inferred DO of about 5.7 mg/L, and streams with TN > 1.12 mg TN/L had a mean inferred DO of just under 5.0 mg/L. None of the reference sites had a mean TN concentration greater than 1.12 mg TN/L nor did any have an inferred DO concentration < 5 mg/L (Figure 3.2). To assure general protection of all regional streams, including those with DO limits of 5 mg/L, the TN concentration at the observed changepoint (1.12 mg TN/L) appears to be a reasonable harm-to-use threshold. The concentration of 1.12 mg TN/L can be considered ecologically meaningful because (1) the reference sites fall to the left of this concentration and (2) it separates the dataset into streams that, on average, always meet the higher regional DO standard (5 mg/L) from streams that do not. It also makes sense that nitrogen correlates well to the diatom OTI as these streams are likely nitrogen limited\textsuperscript{41}.

4.5 Conclusion

Stream-bottom scour and TSS are apparently among the most important variables influencing aquatic plant populations in streams of the Northwestern Glaciated Plains ecoregion of Montana. In spite of the overarching effect of TSS/bottom scour in the streams, we were able to detect the influence of stream nutrient concentration gradients on a diatom metric (diatom Oxygen
Tolerance Index, or OTI; van Dam et al.\textsuperscript{115}). The usefulness of the metric for identifying a harm-to-beneficial use threshold was evaluated by comparing it to five specified conditions that assured it responded to the pollutant of concern (nutrients) and that it was fairly insensitive to other environmental parameters (physical and biological) operating simultaneously in the streams. The diatom OTI passed all 5 conditions, including a close examination of the metric’s ability to discern DO minima caused by nutrient-only \textit{vs.} DO minima caused by nutrient + organic pollution. The metric was significantly correlated to TN concentrations, and we were able to use reference sites, changepoint analysis, and stream DO standards\textsuperscript{8} to identify a nutrient concentration that is ecologically meaningful and should protect regional aquatic life and fish. The final TN nutrient criteria recommended from the analysis is 1.12 mg TN/L.

**Acknowledgements**

Lei Zheng and Jeroen Gerritsen of Tetra Tech, Inc. provided analysis, input, and review on earlier versions of some of the work presented in this appendix. We wish to thank them for this. Thanks to Loren Bahls, who identified the diatoms and provided valuable input on diatom ecology and sampling protocols. Rosie Sada de Suplee of the Montana Department of Environmental Quality assisted with field reconnaissance early in the project and supported the project throughout. We wish to thank John Lhotak, Marianne Zugel, Mike Baker and Pedro Marques of the University of Montana for undertaking field work. Clint Boschen of TetraTech, Inc. assisted with GIS basin-delineation work and analysis. Finally, we would like to thank the private landowners and the peoples of the Fort Peck Reservation who granted us permission to access streams on their lands. Funding was provided by the U.S. Environmental Protection Agency, the Montana Department of Environmental Quality and the U.S. Department of Interior Bureau of Land Management.
Reference List


   Ref Type: Report


   Ref Type: Report


54. Montana Department of Environmental Quality. 2006 Integrated 305(b)/303(d) Water Quality Impairment List and Reports. 2006. Helena, MT, Montana Department of Environmental Quality. Ref Type: Report


Ref Type: Report


Ref Type: Report

Ref Type: Report


Ref Type: In Press


Ref Type: Report


Ref Type: Report

Ref Type: Unpublished Work

Ref Type: Report


Ref Type: Report


Environmental Protection Agency. Ref Type: Report


