Nutrient Reference Condition Identification and Ambient Water Quality Benchmark Development Process

Rivers and Streams within USEPA Region 7
Acknowledgments

The authors gratefully acknowledge the contributions of the following: the USEPA Region 7 RTAG members from the various agencies, tribes, organizations and universities within Iowa, Kansas, Nebraska and Missouri. In addition, the authors recognize invited water quality and biological specialists who attended and contributed at meetings on topics within their range of expertise.

The Regional Technical Assistance Group (RTAG) for the development of nutrient criteria in Environmental Protection Agency (USEPA) Region 7 formed in 1999. The group consisted of state, federal, tribal and academic members. The RTAG’s mission was to develop nutrient benchmarks (surrogate criteria) for midcontinent streams and rivers. The benchmarks developed were designed to protect aquatic life against anthropogenic eutrophication (excess nutrients beyond natural nutrient levels).

The RTAG used USEPA guidance for developing nutrient criteria, and the process consisted of several iterations of data gathering and assessment, stream classification, and statistical analysis and modeling. Stream data gathering and assessment were conducted by the Central Plains Center for BioAssessment (CPCB) with the assistance of RTAG members. Several stream classification methods were pursued including classifying and analyzing midcontinent streams by Level III ecoregions, Strahler stream order, and land use-land cover. RTAG members developed a selection process for identifying reference streams, and reference streams were identified in each state. Statistical analyses were performed on reference stream data to generate reference stream conditions. Reference stream conditions were also modeled using two statistical approaches: percentiles of all stream data and trisection modeling based on chlorophyll-\(a\) levels.

Nutrient benchmarks for streams were determined by the RTAG using a weight-of-evidence approach and operating on group consensus basis. Stream nutrient benchmarks were selected by the RTAG after examination and discussion of stressor and response values derived the analysis of five different assessment approaches: 1) a \textit{priori} determined reference method, 2) quartile method, 3) trisection method to define reference, 4) stressor-response method (e.g. linear and non-linear regressions) and, 5) examination of scientific literature. The RTAG’s final benchmark numbers were developed for the entire Region 7 area that includes the states of Kansas (KS), Iowa (IA), Missouri (MO) and Nebraska (NE).

Benchmarks for streams occurring in Region 7 are as follows:

- 900 \(\mu\)g/L for total nitrogen (TN)
- 75 \(\mu\)g/L for total phosphorus (TP)
- 8 \(\mu\)g/L for sestonic chlorophyll-\(a\) (Chla)
- 40 mg/M\(^2\) for benthic chlorophyll-\(a\) (Chla)

USEPA gratefully acknowledges the contributions of the following: primary authors Walter K. Dodds from Kansas State University, Donald Huggins and Debbie Baker from University of Kansas, Gary Welker from USEPA Region 7 and Regional Technical Advisory Group (RTAG) members from the various agencies, tribes, organizations and universities within Iowa, Kansas, Missouri, and Nebraska. John Bender (Nebraska Department of Environmental Quality), John Olson (Iowa Department of Natural Resources), Mark Osborne (Missouri Department of Natural Resources), Ed Carney (Kansas Department of Health and Environment), Dr. Walter Dodds (Kansas State University), Dr. Don Huggins (Kansas Biological Survey), Dr. John Downing (Iowa State University), Dr. John Jones (University of...
Missouri, Columbia), Dr. John Holz (University of Nebraska, Lincoln) and Dr. Gary Welker (USEPA) were integral members of the RTAG, and their collective knowledge and insight were primary drivers in the development of this document and resulting regional benchmarks. The significant duration in completing this document and making it available should not be taken as a lack of interest or commitment on the part of RTAG members. The ability to bring together so many experts and develop a document based on near total consensus was no small matter, and crafting a final version fell into the hands of a few who committed much time in seeing this document to completion. Dr. Huggins, Debbie Baker (CPCB), and Gary Welker, along with several other USEPA Region 7 staff, crafted the completed version of this RTAG document. While every effort was made to capture and relate all findings and decisions of the RTAG, any omissions and commissions of RTAG results and opinions rest with this last group of individuals.
# Table of Contents

Introduction ........................................... 7  
Overview of the Nutrient Problem ................. 13  
Geographic Area Covered by this Document .... 15  
Data Collection for Streams and Rivers .......... 21  
Assessment of Classification Factors .......... 25  
Statistical Analyses .................................. 29  
Reference Condition Determination ............. 41  
Framework for Establishing Nutrient Benchmarks . 51
Introduction

Background

Nutrients are essential to sustain life and fuel growth. However, in life, too much of a good thing can lead to deleterious effects. In general, excessive amounts of nutrients lead to increased cyanobacteria and algal production, which results in increased availability of organic carbon within an ecosystem. This process is known as eutrophication (Bricker et al. 1999). Nutrients in excess of natural conditions, termed cultural eutrophication, impair aquatic life and can lead to harmful human health effects. Results of excess nutrient levels include production of phytoplankton blooms (cyanobacteria, eukaryotic algae) leading to decreased oxygen concentrations, shifts and loss of aquatic species including fish (e.g., Leach et al. 1977; Eminson and Phillips, 1978; Persson et al. 1991; Schupp and Wilson, 1993; Egertson and Downing, 2004) and excessive macrophyte growth (Daldorph and Thomas, 1978). Further consequences of high nutrient levels are decreases in water clarity (i.e., murky water), drinking water taste and odor problems, and human health effects from the production of cyanobacteria (‘blue green algae’) and their resultant toxins (see review by Chorus et al. 2000).

Nutrients not only affect autotrophic state of flowing waters, but can also alter heterotrophic state (Dodds, 2006). Research in this area is sparse, but researchers have demonstrated that nutrient enrichment of forested streams can influence litter decomposition (Abelho and Graca, 2006), secondary production of invertebrates (Cross et al. 2006), and production of vertebrates (salamanders) that depend upon streams (Johnson and Wallace, 2005; Johnson et al. 2006). Thus, protecting biotic integrity likely will require nutrient control even in systems where autotrophic processes are not dominant (e.g., turbid rivers or streams with a dense canopy cover).
Biotic integrity can also be influenced by the toxic effects of nitrate (Smith et al. 2006). Invertebrate biodiversity is negatively correlated with the nutrient content of rivers and streams (Wang et al. 2007). Some species of unwanted cyanobacteria can be stimulated by nutrients in streams (Perona and Mateo, 2006), and some species of stream cyanobacteria produce toxic microcysts (Aboal, et al.; 2005, Makarewicz et al.; 2009). Research is still needed to directly link eutrophication in rivers and streams to algal toxin production.

Sources of excess nutrients include agricultural runoff, municipal wastewater, urban runoff and atmospheric deposition (USGS, 1999; Mueller and Spahr, 2006). Impairment from excess nutrients in lakes, streams and wetlands has been documented in virtually all 50 States. “Dead zones” in coastal waters resulting from cultural eutrophication have been documented as well. The Gulf of Mexico hypoxic (i.e., low oxygen) zone, an area along the Louisiana-Texas coast, is believed to be the resultant impairment from nutrient contributions in the Mississippi, Missouri, and Ohio River watersheds.

**Nutrient Benchmark Development Process for Streams**

On the 25th anniversary of the 1972 Clean Water Act, former Vice President Al Gore called for the development of an action plan that would fulfill the original “fishable and swimmable” waters goal of the Act. The result of the call to action was the *Clean Water Action Plan* (February, 1998) which provided a “blueprint for restoring and protecting the nation’s waters” by building upon past water quality accomplishments and proposed new challenges for the protection of the nation’s waters. One of the challenges proposed was the reduction of nutrient over-enrichment, and the Environmental Protection Agency (USEPA) was called upon to develop numerical criteria—acceptable levels of nutrients (i.e., nitrogen and phosphorus) in water. Nutrient criteria would be different than typical water quality criteria and would be a “menu” of different numeric values based on waterbody types (i.e., stream, lake, wetland) and ecoregional, physiographic or other spatial classifications.

In June of 1998, the *National Strategy for the Development of Regional Nutrient Criteria* was produced by USEPA’s Office of Water and provided an approach for both: 1) assessing nutrient information and 2) working with States and Tribes in the development of protective nutrient criteria. Key elements of the strategy included: 1) taking a geographic and waterbody approach; 2) developing technical guidance; 3) using regional nutrient teams; and 4) developing criteria by States and Tribes.
Rather than develop nationwide criteria for nutrients, the national nutrient strategy called for the development of criteria based on regional geographic basis as defined by geology, soils, topology, vegetation and climatic conditions. One suggested geographic approach was the “ecoregion” framework (Figure 1), (Omernik, 1987) developed by James Omernik and used as the basis for his later development of nutrient ecoregions (Omernik, 2000), illustrated above (Figure 2). The size of ecoregions can vary from watershed size to continental in scale. For the purposes of developing nutrient criteria, the scale or size of the ecoregions should be dictated by regional nutrient conditions and availability of data. In addition to Omernik’s ecoregion work, other geographic regions and ecoregions have been developed by a number of researchers (e.g., Kuchler, 1964; Bailey, 1995; Maxwell et al. 1995; Abel et al. 2000). While a number of regionalization approaches are available, Omernik’s ecoregions are often suggested in approaches proposed for the development of nutrient criteria (Rohm et al. 2002; Dodds and Oakes, 2004; TDEC, 2004; Stoddard, 2005; VWRRC, 2005).

A major element of the national nutrient strategy was to develop nutrient criteria by waterbody type (i.e., lakes, streams, and wetlands). The focus of this document is on the development of nutrient benchmarks for streams. The technical advisory groups will determine or adapt a classification of streams for criteria development based, in whole or in part, on observed relevant nutrient relationships found for stream classes. The strategy also calls for the development of nutrient criteria for lakes, wetlands, estuaries, and coastal marine waters and highlights the need to keep in mind the inter-relationship of all waterbody types.

Technical guidance

In December of 2000, the USEPA Office of Water published river and stream nutrient criteria recommendations and associated documents for each of the 14 “nutrient regions” in the continental United States. Seven of the 14 documents pertain to nutrient regions occurring within the geopolitical boundaries of USEPA Region 7 (Figure 2): *Ambient Water Quality Criteria Recommendations Rivers and Streams in Nutrient Ecoregion IV, V, VI, VII, IX, X and XI* (USEPA, 2000c-f, 2001c-e). Nutrient criteria recommendations published in the aforementioned documents are to provide guidance for States and Tribes in developing water quality criteria and to provide benchmarks to USEPA when federal promulgation of nutrient standards is deemed necessary. Other uses of the recommendations cited in the documents are the identification of status and trends and the use of yardsticks or benchmarks for over-enrichment assessment in rivers and streams. In addition to these regional stream documents, USEPA’s Office of Water has also produced a similar series of documents for lakes and reservoirs.

**Regional Technical Advisory Group (RTAG)**

The RTAG (i.e., regional nutrient workgroup) for USEPA Region 7 was first established in 1999. The regional workgroup was coordinated by the Region 7 Regional Nutrient Coordinator (Dr. Gary Welker) and facilitated by the Central Plains Center for BioAssessment (Dr. Don Huggins, Director). The workgroup was composed of individuals representing governmental, tribal and academic institutions and having technical expertise in nutrients and water quality standards. Membership has changed slightly since the formation but has essentially been composed of scientists from: Iowa Department Natural Resources; Kansas Department of Health and Environment; Missouri Department of Natural Resources;
Nebraska Department of Environmental Quality; the Prairie Band of Potawatomi Indians; Iowa State University; University of Kansas; Kansas State University; University of Missouri-Columbia; University of Nebraska-Lincoln; U.S. Geological Survey; U.S. Department of Agriculture; Central Plains Center for BioAssessment, and the U.S. Environmental Protection Agency.

The mission of the workgroup was to “develop scientifically defensible numeric nutrient benchmarks for lakes/reservoirs, streams/rivers and wetlands in the Central Great Plains (Iowa, Kansas, Missouri and Nebraska).” The RTAG’s role was to develop benchmarks for nutrients to aid the States and Tribes in their responsibility to develop nutrient criteria. Workgroup operational ground rules for the development of nutrient benchmarks for rivers and streams were as follows:

- Nutrient benchmarks are to protect rivers, streams, and downstream receiving waters against adverse impacts of cultural eutrophication (excess nutrient levels above natural or minimally impaired conditions).

- Nutrient benchmarks are to be protective of aquatic life. Economics, technology, attainability and social values are not part of the benchmark development process.

- Benchmarks developed by the RTAG are by group consensus and are developed for the purpose of assisting and providing guidance to States and Tribes in the development of their own nutrient criteria.

River and stream variables

USEPA requires that nutrient criteria developed by States and Tribes be composed of both causal and response variables (Grubbs, 2001). Nitrogen and phosphorus have long been known to be primary causes of cultural eutrophication (National Academy of Science, 1969; Smith, 1982; Elser et al. 1990; Correll, 1999; Jeppesen et al. 2000; National Research Council, 2000) and have been selected as the two primary nutrient causal variables. The linkage between causal variables (e.g., TP and TN) and chlorophyll-\(a\) (Chl\(a\)) (response variable), a commonly used indicator of algal biomass, is well-known (USEPA, 2000a, b). Both sestonic and benthic Chl\(a\) could be used as algae response variables. Macroinvertebrate and fish communities can also be examined for possible indirect effects related to ecosystem-level changes from nutrient enrichment (i.e. increases in nitrogen and phosphorus). Macroinvertebrate and fish community metrics quantify changes in community structure and function with changes in trophic state. Other examples of response variables that a state or tribe could develop into criteria are periphyton metrics, biological oxygen demand, and macrophyte metrics. In summary, emphasis should be placed on developing criteria for both causal (TN and TP) and response variables.
Development of criteria by States and Tribes

The Office of Water has produced nutrient criteria recommendations (sometimes referred to as 304(a) criteria), which are based on rather large geographic areas (i.e., nutrient regions). The result of the RTAG effort and this document is the development of “nutrient benchmarks” on a smaller geographic scale. However, the states and authorized tribes are ultimately responsible for developing causal and response nutrient criteria for their state/tribe on a statewide/reservation-wide approach or using subsets of geographic areas (i.e., ecoregions) within their state/tribal borders. Thus the States and Tribes, on a geographic scale, will further refine the nutrient benchmarks developed within this document. It is hoped that the States and Tribes will use this document and the lessons learned from the RTAG’s experiences in the development of their own State/Tribal nutrient criteria. Alternatively, they may chose to adopt the nutrient benchmarks within this document as numeric criteria for streams in their state or tribe.
Overview of the Nutrient Problem

Eutrophication is an established water quality management concept and concern reaching as far back as the 1600’s in America (Capper et al. 1983). However, extensive public recognition of this form of pollution in coastal water bodies is relatively recent. The publication “Eutrophication: Causes, Consequences, and Correctives” (National Academy of Sciences, 1969) is often perceived as the technological beginning of American nutrient pollution awareness. The publication addresses the understanding and abatement of this problem primarily in freshwater lakes and reservoirs. We have since come to better understand the problem in streams, rivers, and estuaries, with a focus on: (1) recommending ways to help watershed managers achieve meaningful reductions in the impacts of nutrient over-enrichment in the near-term and (2) identifying areas where future efforts hold the promise of long-term reductions in nutrient over-enrichment and its effects (Nürnberg, 1996; Smith et al. 1999; Anderson et al. 2002; Smith, 2003).

Eutrophication in rivers and streams has received far less attention than that in lakes. For example, only a few schemes exist to classify trophic state in rivers and streams (Dodds et al. 1998; Dodds, 2006). In contrast, entire books have been written about controlling and classifying trophic state in lakes (e.g., Ryding and Rast, 1989). Thus, most regions are starting at a more fundamental level when working to set nutrient criteria for rivers and streams. Nutrient condition is intimately tied to ecosystem structure and function as well as water quality in rivers and streams (Dodds, 2006). Several challenges exist in defining trophic state for streams and rivers. The most daunting is the non-equilibrium nature of lotic ecosystems that increases variance in the relationship between nutrients in the water column and benthic metabolic activity. Another obstacle is that conditions in rivers and streams reflect processes across a large and heterogeneous landscape. Relationships between nutrients and primary producers, as well as consequences on higher tropic levels, are confounded by...
a number of exogenous and endogenous factors including rapid nutrient spiraling (Mulholland et al. 1995), light limitation due to turbidity and shading (e.g., Lowe et al. 1986; Hill, 1996), frequency and duration of spates (e.g., Briggs 1995; Lohman et al. 1992), habitat heterogeneity (Briggs, 1996; 2000), and duration of nutrient enrichment (Elsdon and Limburg, 2008).

Moreover, it is oftentimes difficult to find reference sites with respect to nutrient conditions in streams. It is even more difficult to find reference sites in rivers where anthropogenic impacts are almost certain to occur at multiple locations within the watershed. Such rivers most often suffer cumulative effects across the whole of their larger drainage areas.

In response to this growing awareness of nutrient impacts, USEPA’s National Nutrient Criteria Program encouraged the development of this technical manual to be used by States and Tribes in the reduction of eutrophication of the nation’s freshwaters. This document concentrates on the point at which aquatic life degradation begins as a result of excessive nutrient concentrations in streams within the USEPA Region 7. These concentrations can be used as benchmarks to evaluate eutrophication and as starting points for nutrient criteria derivation.

![Image of cows grazing in a field by a stream with two people in waders and equipment]
The following section provides a general description of Nutrient Ecoregions and Level III Ecoregions examined in this report, and their geographical boundaries. The boundaries and extent of both the Level III Ecoregions and Nutrient Ecoregions occurring partially or wholly within USEPA Region 7 are shown in Figure 1 and Figure 2 respectively. Portions of seven Nutrient Ecoregions are found in Region 7, all of which have been studied by USEPA to determine Nutrient Ecoregion-level benchmark values for TN, TP and Chla (see USEPA, 2000c-f, 2001c-e). The published benchmark values for the seven Nutrient Ecoregions are used in comparing literature values with RTAG benchmark values.

USEPA Region 7 is composed of fifteen Level III Ecoregions, all of which had stream data. However, the Mississippi Valley Loess Plains contained only one site with one data record. The following are brief descriptions provided by Rohm et al. (2002) of the climate, vegetative cover, topography, and other ecological information pertaining to each of these ecoregions.

Aggregate Nutrient Ecoregion IV: Great Plains Grass and Shrublands.

26. Southwestern Tablelands (Omernik Level III)
Unlike most adjacent Great Plains ecological regions, little of the Southwestern Tablelands is in cropland. Much of this elevated tableland is in sub-humid grassland and semiarid range land. The potential natural vegetation in this region is grama-buffalo grass with some mesquite-buffalo grass in the southeast and shinnery (midgrass prairie with open low and shrubs) along the Canadian River.

28. Flint Hills (Omernik Level III)
The Flint Hills is a region of rolling hills with relatively narrow steep valleys, and is composed of shale and
cherty limestone with rocky soils. In contrast to surrounding ecological regions that are mostly in cropland, most of the Flint Hills region is grazed by beef cattle. The Flint Hills mark the western edge of the tallgrass prairie, and contain the largest remaining intact tallgrass prairie in the Great Plains.

43. Northwestern Great Plains (Omernik Level III)
The Northwestern Great Plains ecoregion encompasses the Missouri Plateau section of the Great Plains. It is a semiarid rolling plain of shale and sandstone punctuated by occasional buttes. Native grasslands, largely replaced on level ground by spring wheat and alfalfa, persist in rangeland areas on broken topography. Agriculture is restricted by the erratic precipitation and limited opportunities for irrigation.

44. Nebraska Sand Hills (Omernik Level III)
The Nebraska Sandhills comprise one of the most distinct and homogenous ecoregions in North America. One of the largest areas of grass-stabilized sand dunes in the world; this region is generally devoid of cropland agriculture and is treeless except for some riparian areas in the north and south. Few streams drain this ecoregion, but large portions of the region contain numerous lakes and wetlands.

Aggregate Nutrient Ecoregion V: South Central Cultivated Great Plains

25. Western High Plains (Omernik Level III)
Higher and drier than the Central Great Plains to the east, and in contrast to the irregular, mostly grassland or grazing land of the Northwestern Great Plains to the north, much of the Western High Plains comprises smooth to slightly irregular plains having a high percentage of cropland. Grama-buffalo grass is the potential natural vegetation in this region, as compared to mostly wheatgrass-needlegrass to the north, Trans-Pecos shrub savanna to the south, and taller grasses to the east. The northern boundary of this ecological region is also the approximate northern limit of winter wheat and sorghum and the southern limit of spring wheat.

27. Central Great Plains (Omernik Level III)
The Central Great Plains is slightly lower, receives more precipitation, and is somewhat more irregular than the Western High Plains to the west. Much of this ecological region is now cropland but was once grassland with scattered low trees and shrubs in the south. The eastern boundary of the region marks the eastern limits of the major winter wheat growing area of the United States.
Aggregate Nutrient Ecoregion VI: Corn Belt and Northern Great Plains

47. Western Corn Belt Plains 
(Omernik Level III)
Once covered with tallgrass prairie, more than 75 percent of the Western Corn Belt Plains is now used for cropland agriculture, and much of the remainder is in forage for livestock. A combination of nearly level to gently rolling glaciated till plains and hilly loess plains, an average annual precipitation of 63-89 cm, which occurs mainly in the growing season, and fertile, warm, moist soils makes this one of the most productive areas of corn and soybeans in the world. Major environmental concerns in the region include surface and groundwater contamination from fertilizer and pesticide applications as well as impacts from concentrated livestock production.
Aggregate Nutrient Ecoregion VII: Mostly Glaciated Dairy Region

52. Driftless Area (Omernik Level III)
The hilly uplands of the Driftless Area easily distinguish it from surrounding ecoregions. Much of the area consists of a deeply dissected, loess-capped, bedrock dominated plateau. The region is also called the Paleozoic Plateau because the landscape’s appearance is a result of erosion through rock strata of the Paleozoic Age. Although there is evidence of glacial drift in the region, the influence of the glacial deposits have done little to affect the landscape compared to the subduing influences in adjacent ecoregions. Livestock and dairy farming are major land uses and have had a major impact on stream quality.

Aggregate Nutrient Ecoregion IX: Southeastern Temperate Forested Plains and Hill

29. Central Oklahoma/Texas Plains (Omernik Level III)
The Central Oklahoma/Texas Plains ecoregion is a transition area between the once prairie, now winter wheat-growing regions to the west and the forested, low mountains of eastern Oklahoma. The region does not possess the arability and suitability for crops such as corn and soybeans, which are common in the Central Irregular Plains to the northeast. Transitional “cross-timbers” (little bluestem grassland with scattered blackjack oak and post oak trees) is the native vegetation, and presently rangeland and pastureland comprise the predominant land cover. Oil extraction has been a major activity in this region for more than eighty years.

40. Central Irregular Plains (Omernik Level III)
The Central Irregular Plains have a mix of land use and are topographically more irregular than the Western Corn Belt Plains (47) to the north, where most of the land is in crops. The region, however, is less irregular and less forest-covered than the ecoregions to the south and east. The potential natural vegetation of this ecological region is a grassland/forest mosaic with wider forested strips along the streams than those found in Ecoregion 47 to the north. The mix of land use activities in the Central Irregular Plains also includes mining operations of high-sulfur bituminous coal. The disturbance of these coal strata in southern Iowa and northern Missouri has degraded water quality and affected aquatic biota.

72. Interior River (Omernik Level III)
The Interior River Valleys and Hills area is made up of many wide, flat-bottomed terraced valleys, forested valley slopes, and dissected glacial till plains. In contrast to the generally rolling to slightly irregular plains in adjacent ecological regions to the north (54), east (55) and west (40, 47), where most of the land is cultivated for corn and soybeans, a little less than half of this area is in cropland, about 30 percent is in pasture, and the remainder is in forest. Bottomland deciduous forests and swamp forests were common on wet lowland sites, with mixed oak and oak-hickory forests on uplands. Paleozoic sedimentary rock is typical, and coal mining occurs in several areas.
Aggregate Nutrient Ecoregion X: Texas-Louisiana Coastal and Mississippi Alluvial Plains

73. Mississippi Alluvial Plain (Omernik Level III)
This riverine ecoregion extends from southern Illinois, at the confluence of the Ohio River with the Mississippi River, south to the Gulf of Mexico. It is mostly a broad, flat alluvial plain with river terraces, swales, and levees providing the main elements of relief. Soils are typically finer-textured and more poorly drained than the upland soils of adjacent Ecoregion 74, although there are some areas of coarser, better-drained soils. Winters are mild and summers are hot, with temperatures and precipitation increasing from north to south. Bottomland deciduous forest vegetation covered the region before much of it was cleared for cultivation. Presently, most of the northern and central parts of the region are in cropland and receive heavy treatments of insecticides and herbicides. Soybeans, cotton, and rice are the major crops.

74. Mississippi Valley Loess Plains (Omernik Level III)
This ecoregion fell into the study area of USEPA Region 7 in 2002 with revisions of the ecoregions boundaries and now lies within Ecoregion 73. This ecoregion stretches from near the Ohio River in western Kentucky southwest to Louisiana. It consists primarily of irregular plains, some gently rolling hills, and, near the Mississippi River, bluffs. Thick loess is one of the distinguishing characteristics. The bluff hills in the western portion contain soils that are deep, steep, silty, and erosive. Flatter topography is found to the east, and streams tend to have less gradient and siltier substrates than in the Southeastern Plains ecoregion (65). Oak-hickory and oak-hickory-pine forests were the natural vegetation. Agriculture is now the dominant land cover in the Kentucky and Tennessee portion of the region, while in Mississippi there is a mosaic of forest and cropland.
Aggregate Nutrient Ecoregion XI: Central and Eastern Forested Uplands

39. Ozark Highlands (Omernik Level III)
The Ozark Highlands ecoregion has a more irregular physiography and is generally more forested than adjacent regions, with the exception of the Boston Mountains (38) to the south. The majority of this dissected limestone plateau is forested; oak forests are predominant, but mixed stands of oak and pine are also common. Karst features, including caves, springs, and spring-fed streams are found throughout the Ozark Highlands. Less than one fourth of the core of this region has been cleared for pasture and cropland, but half or more of the periphery, while not as agricultural as bordering ecological regions, is in cropland and pasture.
The USEPA Region 7 RTAG and outside experts comprised water quality specialists, colleges, water resource managers and scientists, representing a variety of state and tribal agencies and universities, who were selected in accordance with National Nutrient Criteria Strategy Document guidelines (USEPA, 1998). Database compilation and data analysis were performed by the Central Plains Center for BioAssessment (CPCB).

In agreement with the USEPA’s National Strategy, the RTAG initially recommended four primary variables for data collection: total nitrogen (TN), total phosphorus (TP), and chlorophyll-a (Chla, sestonic and benthic). These variables were selected, in part, because they are early indicators of cause-and-effect relationships related to nutrient loadings. In addition, a more recent review of hundreds of publications of nutrient enrichment experiments concluded that the absolute concentrations of nitrogen and phosphorus (TN and TP) were the best predictors of periphyton biomass responses (Keck and Lepori, 2012). Early in the RTAG process it was recognized that more nutrient data needed to be collected and analyzed for reference streams to increase their sample size and allow better characterization of this population. USEPA Region 7 tasked CPCB to sample a select number of reference streams for TN, TP and Chla from both sestonic and benthic algal communities. Spring, summer, and fall samples were collected and analysed for each of three years to enhance the reference stream database. Auxiliary data (e.g., ecoregion, stream order, turbidity, biological information) were also collected to link trophic state to other potential abiotic drivers and biotic responses.

Data Sources

CPCB compiled available water quality data for streams and rivers in Iowa, Kansas, Missouri, and Nebraska (Figures 3–4). These data were collected between 1965 and 2003 by a variety of agencies and individuals using
established internal quality assurance procedures (Table 1). All data were ultimately combined into a single Microsoft Access® relational database and checked for accuracy and quality.

Though the data came from reliable sources, the methods of recording the data, units of measure, common and specific name, and resolution of geospatial data sometimes varied from agency to agency. Therefore, CPCB identified differences in data attributes and reporting, and standardized all information within the RTAG databases. Sites that were geospatially similar but had differing information (e.g., fish, water quality, macroinvertebrate) were linked by a proximity rule. All sites along the same stream channel that were within 2 km of each other were given the same numeric site code, if there were no tributaries, wastewater treatment plants, or other conditions between them that might greatly affect their biological and physiochemical similarity.

The Primary Database—Raw Data

The chemistry database used in developing this report was created from the sources detailed above. In the chemistry database, each record represented a single sampling event, and its relevant source information

<table>
<thead>
<tr>
<th>Dataset name</th>
<th>Agency</th>
<th>Project</th>
<th>State</th>
<th>Years</th>
<th>Sites</th>
<th>TN</th>
<th>TP</th>
<th>Chla</th>
<th>Fish</th>
<th>Inv</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPCB DO</td>
<td>CPCB</td>
<td>DO flux</td>
<td>I,K,N</td>
<td>1999-2002</td>
<td>52 177</td>
<td>177</td>
<td>177</td>
<td>0</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>CPCB nutrient</td>
<td>CPCB</td>
<td>reference streams</td>
<td>I,K,M,N</td>
<td>2000-2002</td>
<td>106 726</td>
<td>726</td>
<td>726</td>
<td>0</td>
<td>348</td>
<td></td>
</tr>
<tr>
<td>CPCB TMDL Streams</td>
<td>CPCB</td>
<td>impaired streams</td>
<td>K</td>
<td>1999-2001</td>
<td>42 198</td>
<td>198</td>
<td>198</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>CPCB model</td>
<td>CPCB</td>
<td>KS lake tributaries</td>
<td>K</td>
<td>2001-2002</td>
<td>22 243</td>
<td>243</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IA STORET</td>
<td>USEPA</td>
<td>various</td>
<td>I</td>
<td>1990-1999</td>
<td>268 yes yes yes</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IADNR</td>
<td>IDNR</td>
<td>IDNR</td>
<td>I</td>
<td>1994-2001</td>
<td>207 286</td>
<td>285</td>
<td>0</td>
<td>334</td>
<td>328</td>
<td></td>
</tr>
<tr>
<td>IAREMAP2002</td>
<td>IDNR</td>
<td>IDNR</td>
<td>I</td>
<td>2002</td>
<td>65 176</td>
<td>176</td>
<td>176</td>
<td>885</td>
<td>3984</td>
<td></td>
</tr>
<tr>
<td>IUniv</td>
<td>Univ. IA Hyg. Lab</td>
<td>IDNR</td>
<td>I</td>
<td>1999-2003</td>
<td>99 4104</td>
<td>4388</td>
<td>3856</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>KDHE</td>
<td>KDHE</td>
<td>KDHE</td>
<td>K</td>
<td>1999-2003</td>
<td>318 4422</td>
<td>5822</td>
<td>98</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KDHEnet</td>
<td>KDHE</td>
<td>KDHE</td>
<td>K</td>
<td>1999-2002</td>
<td>20 0</td>
<td>0</td>
<td>0</td>
<td>131</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KS STORET</td>
<td>USEPA</td>
<td>various</td>
<td>K</td>
<td>1990-1997</td>
<td>491 900</td>
<td>11075</td>
<td>727</td>
<td>304</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KDWP</td>
<td>KDWP</td>
<td>KDWP w/REMAP2</td>
<td>K</td>
<td>1994-2001</td>
<td>~399 562</td>
<td>566</td>
<td>0</td>
<td>727</td>
<td>304</td>
<td></td>
</tr>
<tr>
<td>MDNR James</td>
<td>SMSU</td>
<td>MDNR James R.study</td>
<td>M</td>
<td>2001-2002</td>
<td>11 341</td>
<td>334</td>
<td>373</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MDNR streams</td>
<td>MDNR</td>
<td>Ali MDNR data</td>
<td>M</td>
<td>1965-2002</td>
<td>618 9177</td>
<td>13065</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MO STORET</td>
<td>USEPA</td>
<td>various</td>
<td>M</td>
<td>1990-1997</td>
<td>177 1141</td>
<td>1700</td>
<td>257</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MO Jones</td>
<td>UMC</td>
<td>Jack Jones, UMC</td>
<td>M</td>
<td>1985-1986</td>
<td>16 577</td>
<td>579</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOREMAP2</td>
<td>MDOC</td>
<td>REMAP 2002</td>
<td>M</td>
<td>2002</td>
<td>83 0</td>
<td>0</td>
<td>0</td>
<td>83 sites</td>
<td>74 sites</td>
<td></td>
</tr>
<tr>
<td>MDNR ref</td>
<td>MDNR</td>
<td>REMAP 2002</td>
<td>M</td>
<td>2002</td>
<td>61 0</td>
<td>0</td>
<td>0</td>
<td>61 sites</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDEQ2002-2003</td>
<td>NDEQ</td>
<td>N</td>
<td>2002-2003</td>
<td>153 3335</td>
<td>3335</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDEQpre2002</td>
<td>NDEQ</td>
<td>NDEQ</td>
<td>N</td>
<td>1983-2001</td>
<td>194 198</td>
<td>159</td>
<td>0</td>
<td>205 sites</td>
<td>120 sites</td>
<td></td>
</tr>
<tr>
<td>NE EPA Storet</td>
<td>USEPA</td>
<td>various</td>
<td>N</td>
<td>1990-1998</td>
<td>153 1164</td>
<td>1789</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Datasets comprising the USEPA Region 7 stream database analyzed in this document: states (I = Iowa, K = Kansas, M = Missouri, N = Nebraska); number of sample sites; and records of total nitrogen (TN), total phosphorus (TP), chlorophyll-a (Chla), fish, and macroinvertebrates (Inv).
(agency, sampling date) was preserved. As noted previously, the RTAG initially identified five parameters as relevant to Nutrient Criteria development: total nitrogen (TN), total phosphorus (TP), sestonic chlorophyll-\(\text{a}\) (Chl\(\text{a}\)), benthic chlorophyll-\(\text{a}\) (natural substrates), and turbidity. The RTAG also identified two ratios as being of interest, TN:TP and Chl\(\text{a}\)::TP, which were calculated for each sampling event. This chemistry database contained raw sample records of 54,393 sampling events from 1965 to 2003, at 2,400 waterbodies. The number of sampling events per waterbody ranged from 1 to 452. Each record in the database did not necessarily include values for all five parameters. However, most of the records in the database contained some information relevant to Nutrient Criteria development: 32,364 TN values; 51,176 TP values; 8,007 sestonic Chl\(\text{a}\) values; 1,203 benthic Chl\(\text{a}\) values, and 19,087 turbidity values.

The RTAG also directed CPCB to assemble a database of fish collections (2,369 sampling events, 1,325 sites, 1984 to 2003, >200 taxa) and a database of macroinvertebrate collections (1,874 sampling events, 1,151 sites, 1984 to 2003, >1,200 taxa) in USEPA Region 7 (Table 1) as potential indicators of biotic integrity. The raw specimen data was distilled into several metrics, then the metrics for each biological sampling event were paired with chemistry data collected \(\leq 30\) days prior to the biological data. Because so few biological events co-occurred exactly with water quality sampling dates a series of water quality sampling periods was created (i.e. \(\leq 30\), \(\leq 60\), \(\leq 90\), \(\leq 120\) days) and correlated with fish and macroinvertebrate metrics using both Spearman’s rank correlation and Pearson’s correlation (NCSS, 2004). After examination of these correlations, the RTAG concluded that the “best” sampling window was the \(\leq 30\)-day sampling window, given that this window consistently produced the higher correlation coefficients between water quality variables (e.g. TN, TP, turbidity, chlorophyll) and biotic metrics. Metrics calculated were taxa or species richness and percent sensitive taxa for both fish and macroinvertebrate. Also for the fish we calculated Simpson’s Diversity Index (Simpson’s D), and for the macroinvertebrates we calculated percent EPT (Ephemeroptera, Plecoptera and Trichoptera). Using the 30-day window, we paired 1,179 chemistry sampling events with the fish samples, and 507 chemistry sampling events with macroinvertebrate samples.
Macroinvertebrate indices calculations

**Total Taxa Richness**: Count of all taxa found at that site on that date.

**Percent EPT**: Count of all Ephemeroptera, Plecoptera and Trichoptera (EPT) taxa found at that site on that date divided by total taxa richness.

**Percent Sensitive**: Count of all sensitive taxa found at that site on that date divided by total taxa richness. Sensitivity was assigned based on values taken from Regional Tolerance Values in Barbour *et al.* (1999). In that document, for five geographic regions taxa were assigned tolerance values on a scale from 0 (extremely sensitive or not tolerant) to 10 (tolerant). We averaged the literature scores for each taxon and the divided each mean taxa tolerance value by three (3) to reduce the scaling back to three tolerance classes as used in the tolerance scheme for fishes (see below). Therefore the re-scaling of the macroinvertebrate tolerance scores produced the following tolerance scheme: ≤ 3.67 indicated sensitive taxa, the intermediate class was 3.68 to 7.34 and taxa having adjusted tolerances scores greater than 7.35 were considered to be pollution-tolerant taxa.

Fish indices calculations

**Total Taxa Richness**: Count of all taxa found at that site on that date.

**Simpson’s D**: Measures the probability that two individuals randomly selected from a sample will belong to the same species, 0 represents infinite diversity and 1 represents no diversity. The formula is $D = \frac{\sum n(n-1)}{N(N-1)}$ where $n =$ the total number of organisms of a particular species and $N =$ the total number of organisms of all species.

**Percent Sensitive**: Count of all sensitive taxa found at that site on that date divided by total taxa richness. Taxa were marked as sensitive based on two documents that list fish sensitivity values. The first list was Tolerance and Trophic Guilds of Selected Fish Species in Barbour *et al.* (1999). In this document, taxa were assigned tolerance values of I = intolerant (sensitive), M = intermediate, or T = tolerant. The second list was the Autecology table developed by Dave Peck of USEPA’s Western Ecology Division Laboratory in Corvallis, Ore. Peck compiled information for this table on the Autecology of North American fishes for use in EPA’s EMAP program studies. If the tolerance values for the two lists differed for a taxon, the more sensitive category or the category for the “corn belt” region was selected for use in our analyses.
The RTAG identified five factors as potentially important classification variables for use in the initial data analyses. The first factor attempts to describe similar geophysical regions that have similar potential nutrient and biological features. The second factor addresses temporal classes (e.g., sampling season) and their potential effects on benchmark values. The other three factors were quantitative measures of morphometry and hydrology (Strahler stream order, watershed size, and precipitation). All factors and their relationships with other variables were analyzed using both the primary database and the medians database. Details of these analyses are described in the Statistical Analyses section. Other classification factors (e.g., discharge, stream gradient, dominant substrate, geomorphic stream type) were considered by the RTAG, but these data were seldom known for streams and sites in the RTAG database, a limitation which prevented exploring their usefulness.

**Level III Ecoregion (15 categories):** Fifteen Level III Ecoregions (Omernik, 1987) are found in USEPA Region 7. However, only fourteen regions had sufficient data for analyses. River or stream sampling sites were classified according to USEPA-revised ecoregion boundaries (Chapman, et al. 2001; 2002). These Ecoregions may be reclassified if needed into seven Nutrient Regions (USEPA, 1998), but this regional classification scheme was not used in our analyses.

**Sampling Season (all dates, four seasons, monthly, growing season):** The use of individual sampling date data was dismissed early in the RTAG process as the RTAG could not envision a use for the evaluation of daily effects in establishing benchmarks and because sampling events for stressor and receptor indicators were almost never coordinated among streams, watersheds, ecoregions or states. Thus, analysis of stressor data and receptor data by stream population (e.g., large watershed, ecoregion, Region 7) by sample date was considered impractical because
of the highly variable sample size when testing for small-scale temporal effects (i.e. datewise). Assessment of temporal effects by month, season or even growing season were thought to be more feasible when generating benchmark values that might have practical application (Figure 5). Four “seasonal periods” were created by coding sampling dates: Winter (22 Dec-21 Mar), Spring (22 Mar-21 Jun), Summer (22 Jun-21 Sep), and Fall (22 Sep-21 Dec). The “growing season” within Region 7 was estimated to start around the third week of March and end the second week of September. Upon initial evaluation of the raw data, the RTAG decided not to classify stressor data into any temporal groupings. Instead, the RTAG used all available TP and TN data in determining the overall stressor concentrations of streams and rivers in the medians database. These descriptive statistics were used to calculate various stressor properties of the total stream and river population as well as reference groups defined by various approaches (e.g. tri-section).

Most RTAG attempts to discern temporal differences across additional classification categories (e.g. ecoregions, watershed size) were not practical because of the sometimes limited and often temporally clumped nature of the samples within other categories. Additionally, the defining of months or seasons by specific dates or periods assumed that the climate conditions within Region 7 were fairly homogeneous from north to south and east to west, which they are not. Conversely, only by pooling the temporal data for streams could the RTAG gain a reasonable sample size for analysis and assessment of the other identified classification factors.

Use of temporally related variables of interest when examining stressor/receptor relationships was also discussed at length by the RTAG. As previously stated,
seldom were stressor/receptor data obtained on the same day or week or even at precisely the same site in the system. However, properly linking receptors to current or prior existing stressor concentrations was recognized as important in attempting to identify stressor/receptor relationships. This was overcome by examining and creating a 30-day sampling window that linked these two groups of variables (see Primary Database—Raw Data Section). Receptor groups examined were fish, macroinvertebrates, and algae (i.e. chlorophyll \( a \) concentrations), all of which were most often sampled in a more restricted portion of the annual cycle than stressor data. Examination of the monthly distribution of receptor samples showed that, in most instances, these data were confined to the warmer months and most occurred within the proposed “growing season” (Figures 6 and 7). There is some regional evidence that the macroinvertebrate community structure exhibits temporal differences, with summer and autumn clusters being more similar to each other than the winter cluster (late December to mid-April) (Kosnicki and Sites 2011). Nearly all of our paired macroinvertebrate/chemistry samples occurred within the summer/autumn period of this study.

**Strahler Stream Order (orders 1–7):** Strahler Stream order was calculated for 1,239 stream sites. Consistent identification of the stream order of the sampling sites was eventually determined to be unfeasible because of the variable resolution among base maps used for this purpose. The USGS’s National Hydrography Dataset (http://nhd.usgs.gov/chapter1/index.html#_Toc474479766) was one of the databases assessed to see if stream order could be accurately and consistently determined throughout USEPA Region 7. Within the NHD, inconsistencies in mapped stream densities were identified both within and between Region 7 states, which forced the RTAG to look at other stream classification variables such as watershed size.
Watershed Size (size in hectares, four categories): Watershed size was calculated for drainages to 2,149 of the stream sites in the database. A synthetic drainage network was first developed using digital elevation maps (DEMs) and delineation algorithms developed by staff of the Kansas Applied Remote Sensing Program of KBS. Two watershed size variables were used in subsequent evaluation of the effects of watershed size on nutrient and biological variables of interest. Both a continuous variable (size in hectares) and categorical variable (size classes) were examined as potential watershed classification variables. Four watershed size classes were selected: Class 1, <3,200 hectares (<12.2 mi$^2$); Class 2, 3,200–32,000 ha (12.3–123.3 mi$^2$); Class 3, 32,000–320,000 ha (123.4–1,233.5 mi$^2$); and Class 4, >320,000 ha (>1,233.5 mi$^2$).

Potential Area Discharge (PAD): Because broad precipitation differences are common among various geographic regions in USEPA Region 7, the RTAG attempted to identify a surrogate for runoff to watershed area. PAD represents the watershed area weighted by mean precipitation. It was calculated by first determining the 15-year (1989-2003) precipitation average for each watershed pixel in an ArcView GIS coverage using estimates derived from PRISM (Parameter-elevation Regressions on Independent Slopes Model, www.prism.oregonstate.edu). Because some of the watersheds in the database were fairly small, the 4 km PRISM data was resampled into 1 km pixel estimates. The mean precipitation for each watershed (to the nearest 0.001 mm) was calculated from the pixel data that represented the watershed. The mean watershed precipitation estimate (cm) was then multiplied by the total watershed area (M$^2$) and the resulting weighted area divided by 10,000 to get a hectare equivalent. This watershed variable was dubbed potential area discharge (PAD) and was considered a very rough approximation of potential watershed runoff (discharge) under average annual rainfall conditions.
Effect of Sample Size on Parameter Means and Spatial Distribution of Sites

The number of samples per site ranged from 1 to 449 depending on parameter of interest. The number of sites and samples for each site was very limited for both sestonic and benthic Chl, which were of greatest concern as they represented possible response variables. The use of site data with a large number of samples was previously recommended by the RTAG so that the calculated median would better reflect the longer-term temporal conditions associated with each stream parameter of interest at a particular site.

The instream variability of algal biomass (i.e., Chl) was anticipated to be high because of algal growth dynamics, variable light conditions, effects of stream flow and other factors unrelated to nutrient concentrations. Similar variability might also be anticipated with nutrient values and other stream chemistry as these parameters would also change because of seasonality and flow conditions (i.e. run off conditions, base flow conditions). Therefore, site means or medians would offer a better estimate of long-term parameter levels associated with a particular site or stream. However, the selected use of sites with higher numbers of temporal samples reduces the number of sites used to estimate the parameter values for a particular stream class (e.g., stream order) or ecoregion or other geographic region. In order to investigate the potential effects of sample size on estimating the population mean and spatial coverage of the data set, we generated a series of plots that show the change in mean values and number of sites (count) used in estimating the mean for both stressors (Figure 8) and response variables of algae (Figure 9). In most instances there was minimal change (≤20%) in population means between all sites and sites with three to seven values, which was deemed a necessity when calculating median values. However, using only sites with three or more measures of a parameter often resulted in a loss of as much as one...
half of the sites in any one analyses of the data relationships, thus compromising some of the robustness of the datasets. Greater changes in the parameter means and greater loss of sites occurs using sites with more than seven samples in the database. Therefore it was determined that using sites with \( \geq 5 \) samples would minimize changes in the population means and sample size while still allowing construction of box plots, five-number summaries (\url{http://en.wikipedia.org/wiki/Box_plot}) used to examine statistical relationships within and between variables of interest.

In this way, sites with enough samples to generate median and other central tendency values could be used to build the recommended medians database. Trimming the medians database to sets having larger numbers of samples per site has limited the number of sites that are available for analysis, and it greatly altered the population means when compared to the means for the original untrimmed dataset.

**The Medians Database**

The previous assessment of relationships between number of sites with data for each variable of interest, sample size categories (e.g. All, \( \geq 3 \), \( \geq 16 \)) and resulting mean values provided the RTAG a method to select a sample size that: 1) allowed the development of a medians database as suggested in the nutrient guidance documents; 2) allowed calculation of box or violin plots (e.g five-number summaries); 3) maximized site numbers for analyses; and 4) limited changes in site means compared to the site mean determined using all data. The RTAG relied on visual interpretation of the line plots (Figures 8 and 9) to select a sample size that met the five-sample requirement of a box plot while still minimizing both the loss of stream sites and any changes in site means. A sample size of five or more appeared to best
fit these criteria and was selected for use in development of the medians database. However, RTAG members expressed concerns regarding the loss of sites with chlorophyll data, especially benthic Chla, if the ≥5 sample size rule was applied to these parameters. The RTAG’s compromise was to retain all sites with ≥3 sestonic or benthic Chla samples because both the ≥3 and ≥5 sample site means were very similar and yet the ≥3 rule preserved many more Chla sites for analyses (see Figure 9). Additionally, the RTAG used all fish and macroinvertebrate data when examining possible causal relationships among TN, TP and Chla parameters. In most cases, the RTAG analyses of fish, macroinvertebrate and nutrient relationships were based on site means for faunal and nutrient variables measured within the 30-day sampling window. The group acknowledged the fact that box or violin plots could not be constructed for some sites because of the limited number of samples.

Overall, the medians database allowed the RTAG to characterize and quantify most sites, nearly all ecoregions and all of Region 7 using parameter medians for TN, TP, and sestonic and benthic Chla. This was consistent with data reduction methods used in the USEPA’s regional guideline documents (USEPA, 2000a-f, 2001b-d, 2002). Other correlation and regression analyses, as well as most descriptive statistics, used either all data values, site means or site medians, depending upon test restrictions.

It should be noted that a waterbody in the medians database may be represented by median TN and TP values that were derived from measurements taken from the same or different sample events. However, median TN:TP ratios were calculated using only temporal paired TN and TP values. Sampling metadata (e.g., the name of data collector, date, and other information about
specific sampling events) were not preserved with the parameter values for each waterbody in the medians database. Thus, the medians database comprised 2,232 sites from approximately 1,900 waterbodies, with each site represented once by the median value of each parameter.

Water Quality Characteristics and Relationships

A number of nutrient and response variables were examined using several statistical and graphical analysis techniques to identify potentially useful physical classification or geographic factors that might facilitate identification of ecologically similar stream groupings. Initial examinations used all stream data available in the regional streams database. Then, three datasets were analyzed to assess possible temporal differences: 1) all data (all dates); 2) data only collected during an a priori-determined plant-growing season; and 3) non-growing season data. The RTAG agreed that the period starting about the third week of March and ending mid-September would encompass the growing season associated with USEPA Region 7 states, as well as many other states within the Central Plains region. A series of one-way ANOVAs (Analysis of Variance) were performed to test for temporal differences within TN, TP, and sestonic and benthic Chl\(\alpha\) data using the raw data (Table 2A). Only TP and sestonic chlorophyll were found to have significant seasonal differences. In the case of TP the percent difference between all dates and both growing and non-growing season was small, indicating that growing season values were higher than both overall and non-growing season values. The same was true for sestonic algal chlorophyll values, but percent differences in the means were larger than TP differences. The significant statistical seasonal differences noted in TP and sestonic chlorophyll means were thought to be of little importance when examining relationships and identifying possible regional benchmarks for these indicators.

It was assumed that both nutrient and response variables would be more likely to display geographic and stream class affiliations during the growing season, when available nutrients might stimulate and support plant growth. At that time, all trophic levels within lotic ecosystems

<table>
<thead>
<tr>
<th>Test Factors</th>
<th>TN (mg/L)</th>
<th>TP (mg/L)</th>
<th>Sestonic Chl(\alpha) (μg/L)</th>
<th>Benthic Chl(\alpha) (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p value</td>
<td>x diff</td>
<td>p value</td>
<td>x diff</td>
<td>p value</td>
</tr>
<tr>
<td>All dates vs. growing season</td>
<td>0.3213</td>
<td>NS</td>
<td>&gt;0.0000</td>
<td>0.0600</td>
</tr>
<tr>
<td>All dates vs. non-growing season</td>
<td>0.1748</td>
<td>NS</td>
<td>&gt;0.0000</td>
<td>-0.088</td>
</tr>
</tbody>
</table>

Table 2A. Results of one-way ANOVA tests (i.e. GLM ANOVA) to examine for possible seasonal effects on selected nutrient variables. All response variables were log transformed (log + 1) but differences between the means (x diff) are in original measurement units. Non significance (NS) in group means was noted when alpha of ≤ 0.05 was exceeded.

<table>
<thead>
<tr>
<th>Test Factor</th>
<th>Index Period</th>
<th>Response Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecoregion</td>
<td>All Dates</td>
<td>TN 0.000 TP 0.000 Sestonic Chl(\alpha) 0.000 Benthic Chl(\alpha) 0.000</td>
</tr>
<tr>
<td>Ecoregion</td>
<td>Grow. Season</td>
<td>0.000 0.000 0.000 0.000 0.000 0.000</td>
</tr>
<tr>
<td>Stream Order</td>
<td>All Dates</td>
<td>0.000 0.001 0.000 0.064 0.000 0.000</td>
</tr>
<tr>
<td>Stream Order</td>
<td>Grow. Season</td>
<td>0.000 0.009 0.000 0.058 0.000 0.000</td>
</tr>
</tbody>
</table>

Table 2B. Results of one-way ANOVA tests (i.e. GLM ANOVA) tests to examine for possible ecoregion and stream order effects were performed using the median values of response variables for streams having 5 or more samples. Non significance (NS) was noted when alpha of ≤ 0.05 was exceeded.

*Benthic chlorophyll-\(\alpha\) estimates were obtained from sampling periphyton attached to natural substrates in the streams

*Turb = turbidity
are most biologically active. However, using only growing season data greatly reduced the number of streams available for analysis. Thus, RTAG first examined analytic differences associated with the use of data collected over a broad temporal range (all the dates) versus growing season data. Many of the response variables were not normally distributed; thus, power transformations were performed to achieve normality in the data. Next, a series of ANOVAs and ANCOVAs (Analysis of Covariance) were conducted on TP, TN, Chl\text{a} and turbidity data to identify potential treatment (i.e. ecoregions, stream order) and temporal effects (i.e., all dates, growing season). Omernik’s level III ecoregions (Chapman et al. 2001; 2002) and Strahler stream orders (Strahler, 1952) were used as geographical and stream classification variables, respectively.

Results of both ANOVA tests indicated that significant ecoregion groupings were common among all nutrient and response variables (Table 2B). Stream order was not a significant covariate except for the ANCOVA test for sestonic chlorophyll-\text{a}, TN:TP and Chl\text{a}:TP ratios. In all cases ANCOVA and ANOVA results were nearly identical both for analyses using all response data (all dates) and only data collected during the growing season dates (mid-March through mid-September). This suggests that the use of all data would be appropriate and would increase both sample size and spatial coverage. Analysis revealed only one statistical difference between all data and growing-season data: the ANOVA results of turbidity difference for stream orders. The overall similarity of these results, summarized in Table 2B, indicate that there is little advantage to using only growing season data. Thus, data from all seasons were retained for all other analyses.
Ecoregional Characteristics

Descriptive statistics of each parameter in each of fourteen ecoregions are summarized in Table 3. In general, few streams in the Central Oklahoma/Texas Plains (COT), Mississippi Alluvial Plain (MAP), Northwestern Glaciated Plains (NGL), Northwestern Great Plains (NGP) and Southwestern Tablelands (ST) ecoregions could be included in the medians database based on the number of chlorophyll samples for these streams. Examination of Table 3 shows that the small stream sample size in ecoregions MAP, NGP and ST for streams with median seston or benthic chlorophyll values prevent inclusion in ANOVA testing for ecoregions or stream order effects. Additionally, streams in ecoregions Central Great Plains (CGP), COT, Interior River Lowlands (IRL) and NGL could not be included in the ANOVA testing for classification effects (i.e. ecoregion and stream order) on benthic chlorophyll-\(\alpha\), again due to the small number of streams available for testing (\(\leq 2\) stream median values). Thus, statistical testing for ecoregions or stream order effects on chlorophyll variables could only be done on 50 to 75 percent of the ecoregions occurring in USEPA Region 7.

Differences due to stream order were found for TP, TN and TN:TP, as well as for sestonic chlorophyll-\(\alpha\) and the ratio of sestonic chlorophyll-\(\alpha\) to TP. Periphyton (i.e., benthic chlorophyll) differences among stream orders were not significant for either the growing season or all samples.

Examination of the Tukey-Kramer multiple-comparison test results (\textit{post hoc} group mean testing from ANOVA tests) and the violin plots for the various nutrient variables suggests that regional differences were not attributed to individual ecoregions, but several different ecoregional groupings (Figures 11-14). Interpretation of these statistical groups is difficult as groups often change memberships in response to the test variable (e.g. TP, TN, sestonic chlorophyll-\(\alpha\)). It is clear that regional differences do occur when considering all streams (i.e. reference and non-reference), but interpretation of the meanings of these groups may be difficult or of limited value. Similar results were noted when the effects of stream order were assessed for the various nutrient variables. Based on \textit{post hoc} multiple-comparison (Tukey-Kramer) test results, there often appeared to be two stream order groups. First through fifth and

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Ecoregion</th>
<th>n</th>
<th>Mean</th>
<th>Minimum</th>
<th>25th quartile</th>
<th>Median</th>
<th>75th quartile</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN (mg/L)</td>
<td>CGP</td>
<td>172</td>
<td>2.1</td>
<td>0.71</td>
<td>1.25</td>
<td>1.78</td>
<td>2.33</td>
<td>16.01</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>CIP</td>
<td>182</td>
<td>1.59</td>
<td>0.02</td>
<td>0.79</td>
<td>1.12</td>
<td>1.57</td>
<td>32.8</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>COT</td>
<td>3</td>
<td>0.3</td>
<td>0.17</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0.4</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>DA</td>
<td>17</td>
<td>4.88</td>
<td>1.2</td>
<td>3.83</td>
<td>5.2</td>
<td>6</td>
<td>9.16</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>FH</td>
<td>60</td>
<td>1.02</td>
<td>0.13</td>
<td>0.53</td>
<td>0.81</td>
<td>1.42</td>
<td>4.37</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>IRL</td>
<td>50</td>
<td>2.44</td>
<td>0.12</td>
<td>1.17</td>
<td>1.79</td>
<td>2.99</td>
<td>9.21</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>MAP</td>
<td>5</td>
<td>0.48</td>
<td>0.19</td>
<td>0.30</td>
<td>0.46</td>
<td>0.68</td>
<td>0.78</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>NGL</td>
<td>12</td>
<td>3.14</td>
<td>0.28</td>
<td>1.04</td>
<td>1.59</td>
<td>4.79</td>
<td>12.34</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>NGP</td>
<td>13</td>
<td>1.19</td>
<td>0.6</td>
<td>0.69</td>
<td>1</td>
<td>1.55</td>
<td>2.64</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>NSH</td>
<td>30</td>
<td>1.46</td>
<td>0.39</td>
<td>0.7</td>
<td>0.99</td>
<td>1.25</td>
<td>6.37</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>OH</td>
<td>220</td>
<td>1.32</td>
<td>0.09</td>
<td>0.39</td>
<td>0.66</td>
<td>1.55</td>
<td>16.05</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>ST</td>
<td>14</td>
<td>1.02</td>
<td>0.34</td>
<td>0.58</td>
<td>0.64</td>
<td>1.49</td>
<td>2.5</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>WCB</td>
<td>237</td>
<td>5.91</td>
<td>0.79</td>
<td>2.78</td>
<td>5.53</td>
<td>8.5</td>
<td>16.34</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>WHP</td>
<td>44</td>
<td>2.21</td>
<td>0.39</td>
<td>0.99</td>
<td>1.61</td>
<td>3.41</td>
<td>5.85</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>CGP</td>
<td>199</td>
<td>0.32</td>
<td>0.04</td>
<td>0.15</td>
<td>0.24</td>
<td>0.37</td>
<td>2</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>CIP</td>
<td>218</td>
<td>0.3</td>
<td>0.02</td>
<td>0.09</td>
<td>0.13</td>
<td>0.2</td>
<td>18.2</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>COT</td>
<td>3</td>
<td>0.04</td>
<td>0.04</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0.04</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>DA</td>
<td>15</td>
<td>0.08</td>
<td>0.03</td>
<td>0.06</td>
<td>0.08</td>
<td>0.1</td>
<td>0.13</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>FH</td>
<td>71</td>
<td>0.13</td>
<td>0.01</td>
<td>0.06</td>
<td>0.1</td>
<td>0.17</td>
<td>0.46</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>IRL</td>
<td>65</td>
<td>0.48</td>
<td>0.02</td>
<td>0.1</td>
<td>0.19</td>
<td>0.3</td>
<td>7.01</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>MAP</td>
<td>6</td>
<td>0.16</td>
<td>0.03</td>
<td>0.11</td>
<td>0.17</td>
<td>0.21</td>
<td>0.25</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>NGL</td>
<td>13</td>
<td>0.14</td>
<td>0.02</td>
<td>0.1</td>
<td>0.14</td>
<td>0.18</td>
<td>0.23</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>NGP</td>
<td>13</td>
<td>0.17</td>
<td>0.04</td>
<td>0.11</td>
<td>0.15</td>
<td>0.2</td>
<td>0.48</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>NSH</td>
<td>30</td>
<td>0.21</td>
<td>0.11</td>
<td>0.14</td>
<td>0.18</td>
<td>0.22</td>
<td>0.52</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>OH</td>
<td>238</td>
<td>0.52</td>
<td>0.02</td>
<td>0.02</td>
<td>0.04</td>
<td>0.14</td>
<td>8.8</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>ST</td>
<td>15</td>
<td>0.07</td>
<td>0.03</td>
<td>0.04</td>
<td>0.06</td>
<td>0.08</td>
<td>0.14</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>WCB</td>
<td>261</td>
<td>0.24</td>
<td>0.03</td>
<td>0.13</td>
<td>0.2</td>
<td>0.29</td>
<td>2.04</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>WHP</td>
<td>46</td>
<td>0.09</td>
<td>0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>0.11</td>
<td>0.63</td>
</tr>
<tr>
<td>Seston Chl (μg/L)</td>
<td>CGP</td>
<td>8</td>
<td>20.79</td>
<td>4.78</td>
<td>8.11</td>
<td>14.8</td>
<td>38.56</td>
<td>42.65</td>
</tr>
<tr>
<td>Seston Chl (μg/L)</td>
<td>CIP</td>
<td>43</td>
<td>10.24</td>
<td>2.1</td>
<td>6.08</td>
<td>9.26</td>
<td>13</td>
<td>24.87</td>
</tr>
<tr>
<td>Seston Chl (μg/L)</td>
<td>COT</td>
<td>2</td>
<td>8.98</td>
<td>5.5</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>12.46</td>
</tr>
<tr>
<td>Seston Chl (μg/L)</td>
<td>DA</td>
<td>10</td>
<td>4.65</td>
<td>0.9</td>
<td>2.48</td>
<td>4.5</td>
<td>7.63</td>
<td>8.57</td>
</tr>
<tr>
<td>Seston Chl (μg/L)</td>
<td>FH</td>
<td>15</td>
<td>14.9</td>
<td>1.2</td>
<td>9.37</td>
<td>13.05</td>
<td>19.1</td>
<td>43.45</td>
</tr>
<tr>
<td>Seston Chl (μg/L)</td>
<td>IRL</td>
<td>9</td>
<td>32.14</td>
<td>2.95</td>
<td>5.07</td>
<td>13.15</td>
<td>52</td>
<td>127</td>
</tr>
<tr>
<td>Seston Chl (μg/L)</td>
<td>MAP</td>
<td>2</td>
<td>3.54</td>
<td>2.58</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>4.5</td>
</tr>
</tbody>
</table>

Table 3: Descriptive statistics by ecoregion of USEPA Region 7, using the medians dataset (from sites with known watershed sizes and five or more samples for the given parameter).
sixth order streams generally formed one group, while a second group was typically composed of sixth and larger streams and rivers. However, benthic algal biomass (i.e., chlorophyll) differences were not found between stream orders. It should be noted that the sample size for benthic chlorophyll was small and limited to only a few ecoregions (see Tables 2 and 3).

Violin plots allowed the RTAG to visualize the similarities and differences among ecoregions based on a number of important characteristics for each parameter of interest (Figure 10). The violin plot combines the basic summary statistics of a box plot (Tukey, 1977) with the visual information provided by a local density estimator. It is a standard box plot surrounded by an outline indicating the data as density estimated by a kernel method (NCSS, 2004). The plot shows the median value with a circle, and indicates the 25th to 75th quartile values using interior lines on either side of the median circle. The resulting figure is then “boxed” by the mirrored density curves for all data used in the analysis. The goal is to not only define the typical box plot measures, but to reveal the distributional structure within a variable. Examination of the density curve shape can be used to distinguish departures from normal distributions. While there are many variations of the box plot, five or more data points are needed to correctly calculate and construct a box plot. Therefore, in this document, a violin plot is provided only when a parameter is recorded for five or more sites represented by individual median values. For example, to generate plots by parameter(s) by ecoregion(s), only those ecoregions having data for ≥ 5 or sites (e.g., streams) which have ≥ 5 measurements of a parameter could be plotted. In representing sites and ecoregions, the use of medians prevents data distortions caused by uneven sampling efforts and yet provides a good estimate of prevalent conditions. For the purposes of this document, outliers are cut off to better show the violin plots.

In general, most of the nutrient variables are right-skewed or “skewed to the right,” indicating that most of these variables have long tails of high values. The few exceptions are the more normal appearing distributions of TN in DA ecoregion and TP in DA and NGL ecoregions. In addition, apparent bimodal distributions occurred in TP within MAP ecoregion, in sestonic Chlα within DA and NSH.

Table 3 cont’d. Descriptive statistics by ecoregion of USEPA Region 7, using the medians dataset (from sites with known watershed sizes and five or more samples for the given parameter).
ecoregions and in benthic Chla in FH, NSH and WHP ecoregions. However, all of these parameters that show bimodal distributions within specific ecoregions were characterized by having low sample sizes (i.e., number of streams). Small sample sizes can cause the density curves to take on a bimodal or wavy appearance since the limited sample size tends to create breaks and dips in the curve-fitting process.

### Stream Size Classification

The systematic determination of stream orders within USEPA Region 7 was attempted then abandoned because of inconsistencies in mapped stream features and variability in the resolution of the source maps. It was clear that not all stream segments, especially small stream reaches (e.g., first order, intermittent stream segments), were mapped on individual USGS 7.5 quad maps, which were the primary maps used in determining stream order. Instead of classifying stream size by stream order, the RTAG investigated the use of watershed size as a surrogate measure of stream size.

### Watershed Size

Watershed sizes were delineated using a synthetic stream system developed from DEMs of the region. The synthetic stream system allowed researchers to use the stream location of each collection site as the watershed outlet and then generate a watershed area capturing all of the upstream area that drains to the collection site. A regional data set of sample sites with both stream order (Strahler, 1964) and watershed size (ha) was generated to examine the potential relation between these two variables. The RTAG checked these data for consistency and quality, and corrected all errors that were apparent in the screening process. Then simple linear and robust regression tests (NCSS, 2004) were
performed on the 1,232 paired watershed size/stream order values (Figure 15). Both the simple and robust linear models produced highly significant models that explain 78 to 86 percent of the variance between these variables, respectively. Based on these results and the inter-regional problems of accurately determining stream order values, we adopted watershed size as a potential stream classification approach.

Next, a series of robust regressions were run between watershed size (i.e. stream classification variable) and a number of stressor (abiotic) and response (biotic) variables (Table 4). These results suggested that about 42 percent of the sestonic chlorophyll concentrations and 22 percent of fish richness were explained by watershed size, with larger watersheds (and streams and rivers) having greater sestonic chlorophyll-\(a\) concentrations and higher fish richness. However, in most instances, there appeared to be little or no relationship between the size of a watershed and either nutrient concentrations or other water quality or biotic variables. While total phosphorus had a significant and positive relationship with watershed size, very little of the variance of TP was explained by this independent variable. It should be noted that the large sample sizes associated with most of variables of interest often cause regressions to meet the conditions of significance (low \(p\) values).

Four watershed size classes were subjectively determined from the distribution of all the watersheds examined in this study (Figure 16) so that the statistical characteristics of nutrients and algal chlorophyll for each group could be calculated. The largest and smallest watersheds for which data were available were 8,159,804 to 136 hectares, respectively (Table 5).

RTAG examined the median and 25\textsuperscript{th} quartile values for nutrients, turbidity, sestonic and benthic chlorophyll-\(a\), and fish macroinvertebrate and richness variables. These values indicated that most watershed classes had similar attributes (Tables 6 and 7). Only one strong trend was noted: sestonic chlorophyll-\(a\) appeared to have increasingly higher median and 25\textsuperscript{th} quartile values as watershed class size increased. Median and 25\textsuperscript{th} quartile values for sestonic chlorophyll in watershed class 4 were more than twice those found for watershed class 3.
To account for the precipitation gradient that occurs from the western to the eastern portion of the Central Plains, the RTAG explored the ways in which PAD (a proxy for potential runoff) relates to discharge, flow, and the biological communities most likely to be affected by flow. A comparison of watershed sizes and their corresponding PAD values in various ecoregions indicated that nearly all ecoregion relationships were statistically significant (p < 0.05 for simple linear regression, GLM model). When we examined the linear relationships between PAD and watershed size, we found that the slopes of these regressions decreased as we moved westward through Region 7 (Figure 17). It was concluded that while the regression slopes did change among ecoregions, the significant positive linear relationships found between PAD and watershed size for all ecoregions were highly explanatory (R² > 0.70), and thus either factor could be used as a stream classification variable. Because watershed size was simpler to calculate and understand as a stream classification variable, the RTAG decided not to pursue the use of PAD as a classification factor in this effort.

However, before PAD was dropped from consideration as a stream classification factor, the relationships between these two classification variables and both macroinvertebrate and fish richness variables were explored using simple least squares regression and LOWESS smoothed trend lines (Figures 18 and 19). The relationships observed between biological variables and both PAD and watershed size were, for all practical purposes, identical. This confirmed that using PAD as a classification variable was unnecessary. These figures also showed that both macroinvertebrate and fish richness tended to peak in watersheds that are from 10^4 to 10^5 hectares in size.
Nutrient Limitation

Limitation by nutrients can be indicated by deviations from Redfield ratios of nitrogen and phosphorus (Dodds et al. 2002). A ratio of 16:1 N:P by moles, or 7:1 by mass, indicates growth. Ratios substantially less than this indicate nitrogen limitation, and substantially greater indicate limitation by phosphorus. Grimm et al. (1981) suggested that an N:P ratio less than 15 indicates that their study streams were nitrogen limited, while N:P ratios greater than 15 were phosphorus limited. Other studies (Grimm and Fisher, 1986; Lohman et al. 1991) also noted that N:P ratios can be used to predict which major nutrient might be limiting to plant growth in streams. For a river community, Schanz and Juon (1983) determined that nitrogen was limiting at N:P < 10, phosphorus was limiting at N:P > 20, and in the inter-range of 10 to 20, neither nutrient could be assumed to be limiting with any level of certainty.

However, others have found N:P ratios to be poor predictors of periphyton biomass (i.e., chlorophyll concentrations). In New Zealand streams, N:P ratios were of limited use in predicting periphyton biomass accrual.
In experimental stream studies, Stelzer and Lamberti (2001) found that early in their experiment, nutrient ratios did not indicate which nutrient (N or P) was limiting; however, higher TN concentrations were consistently associated with higher periphyton biovolume values. The RTAG found that TN was the only independent nutrient variable that explained any reasonable amount of variability in benthic chlorophyll.

About half of the streams examined in the regional database had TN:TP ratios of 10 or less. Most of the ecoregions had N:P ratios close to 7 (from 5.5 to 11.3). The MAP ecoregional values indicated nitrogen limitation (N:P by mass = 2.7), whereas OH (16.5), WCB (29), and WHP (161) had values potentially indicating phosphorus limitation. Because these ratios were determined from all data, including impacted sites, they do not indicate the reference condition. However, most sites seemed to be balanced close to nitrogen and phosphorus limitation, which justifies including both in the nutrient criteria process.

This approach by the RTAG of identifying benchmarks for both TN and TP has been supported by several recent publications, most notably that of Keck and Lepori (2012). Their review of more than 380 nutrient-enrichment experiments suggested that nutrient ratios are only of predictive value at extreme N:P ratios (i.e. <1:1 and >100:1). The RTAG also felt that both nutrients need possible management, for two reasons: (1) the seasonal dynamics of nutrient ratios, and 2) differing tributary non-point source contributions, which may cause potential shifts in ratio values within a single stream network. Shifting temporal and spatial TN:TP ratios within stream and river networks were also reasons given by Wilcox et al. (2007) in seeking control of both macronutrients to limit excessive periphyton growth in some Australian streams.
In order to assess the impacts of human-mediated disturbances, scientists often identify sites that experience relatively minimal levels of impairment and therefore represent “healthy or acceptable” conditions. These reference conditions can then be used as benchmarks for ecosystem health in the development of nutrient criteria (USEPA, 1996; USEPA, 2000a; Stoddard, et al. 2006). USEPA suggested using the following general approaches to assist in identifying and defining reference conditions in streams and rivers:

- Biological survey of sites (determination of reference sites);
- Evaluation of historical data;
- Prediction of expected conditions using models (simulation, statistical, hybrid models);
- Expert opinion/consensus.

However, the question arises as to the meaning of reference condition and the extent of disturbance, or lack of disturbance, it represents. Varying levels of human disturbance found in the environment require the need for a range of reference condition definitions (Hughes, 1995; Bailey et al. 2004; Stoddard et al. 2006). Historical condition refers to the condition of the ecosystem at some point in the past when it was totally undisturbed by human activity. That is, the ecosystem was in absolutely “natural” or pristine condition. Waterbodies in this type of condition are unlikely to be found and are difficult to define, but knowing this condition or approximate condition allows us to better describe all current conditions and the extent of change. The minimally disturbed condition (MDC), or the absence of significant human disturbance, if this can be determined, may serve as a benchmark when comparing other definitions of reference condition, definitions which may change over time in response to changes in climate, land use, and management practices. For example, least disturbed condition describes the “best” condition of water bodies occurring in moderately to heavily altered conditions.
landscapes (e.g., the agroecosystem landscapes of USEPA Region 7). *Best potential condition* (BPC) is defined as the least disturbed ecological conditions achievable with best management practices in place for a time period that yields results. Most states in the Central Plains region refer to the *least disturbed* reference condition due to the extent and nature of modern land use in the plains region. The RTAG examined several different methods of defining a reference condition that represents the least disturbed condition, or what is considered to be the *best potential condition*. This section compares these methods.

The RTAG used several alternative methods to derive possible “reference” or benchmark conditions (e.g., concentrations, values), and these were compared to each other in a weight-of-evidence approach. Some of these methods were literature-derived (back-calculation to ambient nutrient concentration), whereas others were essentially biotic indices (e.g., diversity versus nutrient thresholds). Each of these will be described in the following sections.

**A Priori-Determined Reference Sites**

The Clean Water Act mandates the establishment of narrative biological criteria as part of state water quality standards. When implemented, biological criteria will: (1) expand and improve water quality standards programs; (2) help identify impairment of beneficial uses; and (3) help set program priorities. The USEPA Region 7 Biocriteria Workgroup was formed to develop regionwide guidelines for identifying biological indicators, characterizing biological condition and establishing biological criteria or benchmarks. One outcome of the workgroup’s efforts was a set of guidelines to consider when evaluating reference conditions, the “Core Factors to Consider in the Selection of Reference Condition in Central Plains Streams” (Table 8). The four states in Region 7 submitted lists of stream sites that reflect reference condition, or minimal anthropogenic impact. This list of 308 sites was last updated in 2004.

Nebraska sites were selected by the NDEQ, who evaluated sites sampled for the 1997-2001 Regional Environmental Monitoring and Assessment Program (REMAP). Only the REMAP sites were considered because this program resulted in the best and most complete suite of site data. After choosing sites based on the best habitat scores, NDEQ secondarily used the IBI and ICI scores to verify the best sites. This resulted in 50 reference sites.

Iowa sites were submitted by the Iowa Department of Natural Resources (IDNR), which chose 111 reference sites that are regionally representative and that are least disturbed by human activities (Wilton, 2004). IDNR staff developed guidelines that specify the target number of sites for each ecoregion and the range of stream sizes to be considered for reference site nomination (IDNR, 1992). The population of candidate streams included wadeable streams currently designated for protection under the *Biological Assessment of Iowa’s Wadeable Streams Bioassessment Framework*, which preserves

<table>
<thead>
<tr>
<th>Factors</th>
<th>Primary or secondary evaluator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wastewater treatment plants and other point sources</td>
<td>Primary</td>
</tr>
<tr>
<td>Animal feeding/grazing operations</td>
<td>Primary</td>
</tr>
<tr>
<td>Instream habitat</td>
<td>Primary</td>
</tr>
<tr>
<td>Riparian habitat</td>
<td>Primary</td>
</tr>
<tr>
<td>Land use and land cover—broad scale</td>
<td>Primary</td>
</tr>
<tr>
<td>Land use and land cover—site-specific</td>
<td>Primary</td>
</tr>
<tr>
<td>Physical and chemical parameters</td>
<td>Primary</td>
</tr>
<tr>
<td>Altered hydrologic regime</td>
<td>Primary</td>
</tr>
<tr>
<td>Biological metrics</td>
<td>Secondary/ confirmatory</td>
</tr>
<tr>
<td>Biotic assemblages</td>
<td>Secondary/ confirmatory</td>
</tr>
<tr>
<td>Representativeness</td>
<td>Primary</td>
</tr>
</tbody>
</table>

Table 8. Summary of the factors used in identifying and defining reference sites and conditions within USEPA Region 7 (Biocriteria Workgroup, 2000).
streams for warm water or cold water aquatic life uses. IDNR guidelines excluded both intermittent headwater streams classified as general use and large, non-wadeable interior or border rivers. In reviewing candidate reference sites, IDNR staff considered five major factors: 1) animal feeding operations; 2) channel alterations; 3) land cover/land use; 4) riparian and instream habitat characteristics; and 5) wastewater discharges.

When selecting Kansas sites, Kansas Department of Health & Environment (KDHE) considered many variables, including: water chemistry data, stream flow (trend) data, watershed land-use, municipal and industrial point sources, confined animal feeding operations, impoundments and other channel obstructions/modifications, oil field development activities, and irrigation development activities. In the past few years, KDHE also has considered the extent to which the contemporary biological condition deviates from the historical biological condition, where known (limited primarily to an assessment of fish and shellfish communities).

Missouri Department of Natural Resources (MDNR) used the six-step selection process of Hughes et al. (1986) which provides a flexible and consistent method of evaluating reference suitability (Table 9). Topographic maps, water quality staff at the MDNR and the Missouri Department of Conservation (MDOC), and fisheries management biologists at MDOC were consulted during steps 1, 3, 4, and 5 of the reference stream selection process (Sarver et al. 2002). Water quality violations and fish kill reports were also examined to help in the process. This process resulted in the selection of 72 reference sites in Missouri that were included in this effort.

The population of a priori-selected sites was designated as the “reference” population. Table 11 displays descriptive site statistics for TN, TP, and sestonic and benthic Chl_a. Regionwide, the mean concentration for each indicator was 14 to 30 percent higher than corresponding median values, indicating the data was skewed toward higher values by a few sites with very high values. When examining the relationships between levels of nutrients associated with the reference population (Table 11), it must be remembered that the a priori reference population represents stream/watershed systems that are believed to be minimally impacted by a number of factors (e.g., point source pollutants,
nutrients, altered riparian condition, instream habitat, land use). The a priori reference population may or may not represent sites experiencing minimal nutrient impacts. Similarly, these reference streams do not represent streams with the best biological metrics or biological condition. This is illustrated when we compare the median values for fish and macroinvertebrate richness in the a priori reference population with the same values for the Trisection method, based on the same two endpoints for all sites (Table 14). The very best third of streams, based on fish or macroinvertebrate richness, had median values that were 16 to 26 percent higher, respectively, than the a priori reference streams.

The Quartile Method to Establish a Reference Condition

USEPA’s nutrient technical manuals (USEPA, 2000a, 2000b, 2001b) describe two ways of establishing a reference condition. The first method is to choose the upper quartile value (75th percentile) of the distribution of an a priori reference population of sampling stations (such as those determined in Section 6.1). When reference conditions are not identified, the second method is used. The second method determines the opposite end of the distribution, or the lower quartile value (25th percentile) of the population from all available sampling stations. This method presumes that some of the region’s sampling stations are degraded (Figure 20).

Using the medians dataset, the 25th percentile of the entire stream population was calculated for each parameter (Table 12 and Table 13). This cut-off was later compared to other potential benchmark values. The values also show that this dataset of USEPA Region 7 streams does not fit the theoretical model proposed by the USEPA, since the 25th percentile of the parameters from the entire population is below that of the 75th percentile of the reference population. Histograms of the values further illustrate this point (Figure 21). The model most likely exaggerates what one would find for any set of waterbodies, because it separates the reference population from the entire population. In reality, the reference population is a subset of the entire population and its bell curve should greatly overlap the left side of the bell curve of the entire population.

However, most waterbodies in Region 7 have been already strongly impacted by anthropogenic nutrient loading. In cases where a group of waterbodies already shows evidence of human impact, a percentile other than 25% than can be used in an effort to approximate previous natural conditions (USEPA Nutrient Criteria Guidance Manual, April 2000). The RTAG thus chose to explore a third procedure termed the Trisection method.
The Trisection Method to Establish a Reference Condition

The Trisection method (USEPA, 1998) designates the reference condition for a specific response metric based on the condition of the group of streams that fall in the best third for that metric (Figure 22). This population of reference streams can then be used as a standard of comparison for the entire population of streams. Some methods recommend trimming off the worst 5% of streams before sectioning the population into thirds, but the RTAG ultimately decided not to do this, creating a more conservative estimate.

The RTAG determined the Trisection reference populations using four different biological response variables: fish species richness, macroinvertebrate taxa richness, and benthic and sestonic chlorophyll-\(a\) concentrations. The population of data for each response parameter in the medians dataset was first divided into thirds, and then the waterbodies in the best (i.e., highest values for richness, lowest values for chlorophyll) one-third of each parameter were designated as the “reference” population. The RTAG then examined TP and TN values associated with these reference populations. Median values for both stressor and response parameters were calculated for the reference waterbody population of each response parameter (Table 14). For all but taxa richness, calculations were based on data from sites that had five or more samples collected at each site. Fish and macroinvertebrate richness values for each site were calculated as the mean value for that site, as many sites had fewer than five fish or macroinvertebrate samples. The best fish and macroinvertebrate populations had median richness values of 19 and 66, while median concentration of Chl\(a\) for the reference population based on the lowest one-third sestonic or benthic Chl\(a\) levels were about 11 and two, respectively. It was noted that the reference population based on one biological response variable did not necessarily have the correspondingly best value for

![Figure 21](image_url)

Density traces of non-reference (solid line) and reference (dashed line) streams in USEPA Region 7 for total nitrogen (mg/L), total phosphorus (mg/L) and chlorophyll-\(a\) (\(\mu\)g/L).
other biological parameters (Table 14). For example, the reference population based on fish richness had a median macroinvertebrate richness value that was 35 percent less that the median reference value for macroinvertebrate streams. However, while the reference streams for either sestonic or benthic chlorophyll-\(a\) had low fish richness (median \(\approx 7\)), the median macroinvertebrate richness was quite high (median = 77 to 90). Among the different reference populations for each response parameter, median TN and TP values varied from 5.73 to 0.64 mg/L and 0.07 to 0.12 mg/L, respectively. Reference populations based on macroinvertebrate richness and sestonic Chl\(a\) had the most restrictive median levels for both nutrients, but the number of streams comprising the reference population for sestonic chlorophyll-\(a\) was only 23 (Table 14).

### Regression and Threshold Methods

Potential causal relationships between nutrient stressors and various biotic characteristics of lotic ecosystems in the region were explored with robust regression analysis as a linear modeling technique. This regression technique was used because it is less affected by heteroskedasticity and outliers than ordinary least squares regression (see [http://en.wikipedia.org/wiki/Robust_regression](http://en.wikipedia.org/wiki/Robust_regression); Western, 1995; Rousseeuw and Leroy, 2003). Only two of four robust regression models were found to be statistically significant (alpha \(\leq 0.05\)) when chlorophyll variables were identified as the dependant variables. Sestonic chlorophyll levels were positively related to TP values, indicating that increases in this nutrient contributes to increases in algal biomass. This regression was highly significant and explained 20 percent of the variation in water column chlorophyll values. Conversely, TN was the sole nutrient linked to changes in benthic chlorophyll concentrations (i.e. periphyton). The TN model showed a positive relationship between chlorophyll and TN levels, explaining just over 10 percent of the variation in benthic chlorophyll. In northern Ozark streams in Missouri, Lohman et al. (1992) and Lohman and Jones (1999) found much stronger linear regression relationships (\(R^2 \geq 0.47\)) between benthic algae and both TP and TN summer means, as well as between TP and mean summer sestonic chlorophyll (\(R^2 \geq 0.67\)). More recently, researchers confirmed that in low-level nutrient streams the Ozark Highland ecoregion, fish-, macroinvertebrate- and algal-based biotic indices all showed significant negative relationships with nutrients (Justus et al. 2010). These strong stressor/response relationships in low-nutrient Ozark streams were not found when analysing our regional stream data. This suggests that in most streams in Region 7, linear stressor/response relationships are masked by numerous high stressor values, specifically high nutrient levels, that lie well beyond the stressor/response thresholds (see threshold discussion below). Using data from a large series of temperate streams, Van Niewenhuyse and Jones (1996) found that TP and mean summer benthic chlorophyll values were positively related with simple linear regression, explaining about 67 percent (\(R^2 = 0.67\)) of the variance in chlorophyll concentrations. Identifying strong stream nutrient and plant relationships is often hampered by a number of other stream factors that affect algal biomass accrual and standing crop. One set of factors of concern by the RTAG was hydrologic conditions associated with sampling conditions. Virtually all of the stream data used in these assessments lacked associated discharge and flow measurements. Lack of flow data is particularly concerning because current and prior flow conditions, especially stream velocities, can both increase and decrease production and biomass of periphyton regardless of nutrient conditions (Horner and Welch, 1981; Humphrey and Stevenson, 1992; Biggs, 2000).
Existence of biological thresholds could indicate reference conditions, or conditions under which biological integrity can be compromised. This approach makes the assumption that organisms in rivers and streams have adapted to a range of nutrient conditions. Thus, maximum diversity of organisms can be found within the range of water column nutrient concentrations under which these organisms evolved. If there are thresholds in diversity (e.g., an abrupt change in the relationship between diversity and nutrients) related to an environmental parameter, this indicates environmental conditions outside those typical of a reference condition (see Evans-White et al., 2009). Muradian (2001) provided a comprehensive review of theory and existence of ecological discontinuities and thresholds, and ecological phenomena involving non-linear dynamics. The presence of thresholds or change points might be suspected when existing knowledge indicates probable ecological relationships, and bivariate analysis (e.g. scatterplots, simpler linear regression) reveals discontinuous or non-linear distributions (see Figures 18 and 19). Recently, a number of researchers have identified water quality threshold values for several biological groups, as well as suggested metrics for use in water quality and nutrient management (see review of Evans-White et al., 2013).

Various statistical and mathematical methods have been developed to determine non-linear piecewise or segmented regression and thresholds (e.g. change points, breakpoints) (see Dodds et al., 2010). A simple method is to visually inspect locally weighted regression lines (LOWESS, Cleveland 1979) for regions of abrupt change. Several statistical methods also exist to indicate thresholds, non-linearities or breakpoints in relationships. Breakpoint regression finds the two best lines that fit two-dimensional data. The point of transition from one predictive line to the next indicates a threshold. Two-dimensional Kolomgorov-Smirnov test is a nonparametric method to indicate breakpoints in variance (Garvey et al. 1998). Polynomial regression can indicate whether a non-linear fit explains more variance than a linear fit, but does not establish the breakpoint in the relationship. Piecewise regression has been recognized as a statistical tool for identifying ecological thresholds (Tom and Lesperance, 2003). Breakpoint values for a number of different response variables were determined using a two-segment piecewise, non-linear regression technique (Table 15) given in SigmaPlot 2000 (version 6.00. SPSS, Inc., Chicago, IL). This was done using TP, TN and turbidity as stressor variables. Trimmed versions of both TP and TN medians were also tested against various dependant variables (see foot notes 1-4, Table 15). Trimmed values were visually identified as outliers, and no formal outlier tests were conducted to determine statistical outlier values. Macroinvertebrate richness and ratio of sensitive taxa values from Iowa were dropped from analysis because larvae of the family Chronomidae were not taken to the genus level, as was done for other states, thus altering the taxonomic resolution of the Iowa data.

Later analyses using the software program SegReg (http://www.waterlog.info/) that analyzes for segmented linear regression (i.e., piecewise regression) with one dependent and one or two independent variables gave somewhat different breakpoint results than those listed in Table 15. In SegReg, the significance of the breakpoint is indicated by the 90% confidence area around the breakpoint as shown in the graphical output. When the confidence area remains within the data range, the breakpoint is significant.
The best segmented regression model and breakpoint from this analysis is the model that explains more variability associated with the dependant variable when compared to a simple linear regression of the same dataset. We chose SegReg as the second breakpoint-determining analytical method because SegReg has been widely accepted and used in the natural sciences over the last 25 years. Its recent use in soil science, hydrology, and natural science research is well-documented in the literature (e.g. Wojciechowski et al., 2014; Nordin et al., 2009; Dias et al., 2013; Monteagudo et al., 2012; Ye et al., 2014; Rotvit and Jacobsen, 2013; Minton et al., 2012).

With the exception of macroinvertebrate richness values, the dependent and independent variables used in SegReg modeling were taken from the medians databases. Mean values for macroinvertebrate richness were based only on those stream or river sites that had associated median values for nutrients.

**Figure 23**

Relationships between seasonal mean water column nutrients (total N and total P) and proportion of instances in which seasonal mean and maximum chlorophyll exceed 50, 100 or 150 mg m$^{-3}$. Data from literature sources compiled in Dodds et al. (2002). This compilation previously had incorrect values for data reported by Lohman et al., (1992), however, these incorrect values now have been corrected for use in this figure. Sample size for TN and TP was n = 199 and 250, respectively. (Figure reproduced from Dodds 2006).
We considered the use of medians to describe prevailing nutrient levels to be a conservative approach when compared to the approach of most nutrient stressor/response studies, which use maximum and/or mean values for analysis (e.g., Van Niewenhuysen and Jones, 1996, Dodds et al., 1997; Lohman and Jones, 1999, Briggs 2000; Dodds et al., 2002). In addition to using median values, we tested for outliers using both the outlier labeling with boxplot procedures and the Mahalanobis distance of each point from the variable means (see Hodge and Austin, 2004; Sim et al., 2005 and Hintze, 2013). A small number of statistically significant severe outliers were detected and removed from the analysis; the outlier occurrence was random among the dependent and independent variables.

Although the confidence-area test SegReg makes other types of significance tests unnecessary, we also used ANOVA and applied the F-test to determine significance. The ANOVA procedure assumes that the regression is done of y (dependent variable) on x (independent variable). A number of significant segmented regressions and corresponding breakpoints for TP and TN stressors were identified and are presented in Table 16.

Published Information Related to Nutrient Criteria in USEPA Region 7

Several peer-reviewed publications have estimated nutrient reference levels for ecoregions that occur in USEPA Region 7 (Table 17). Smith et al. (2003) used modeling approaches to estimate reference levels, while Dodds and Oakes (2004) used an analysis of co-variance and extrapolation to remove the influence of human land use. These data are compared to the USEPA method that simply chooses the 25th percentile of all sites. The methods roughly agree, with the largest disparity occurring in the Corn Belt; in this ecoregion, the 25% method yields substantially higher TP and TN estimates, probably because of the absence of extant true reference sites in this ecoregion.

Current literature suggests that ecological thresholds may occur between specific stressors and receptors (e.g., Huggett 2005; Betts et al. 2007, Brenden et al., 2008; Evans-White et al., 2009; Francesco and Denoël, 2009; Miltner 2010; King and Baker, 2010; Caskey et al., 2013; Rotvit and Jacobsen, 2013). A 2002 study

<table>
<thead>
<tr>
<th>Dependant variable</th>
<th>Independent variable</th>
<th>Sample size</th>
<th>P value</th>
<th>R²</th>
<th>Breakpoint for independent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sestonic chlorophyll-a</td>
<td>TP (mg/L)</td>
<td>463</td>
<td>&lt;0.0001</td>
<td>0.11</td>
<td>0.32</td>
</tr>
<tr>
<td>Sensitive fish species</td>
<td>TP (mg/L)</td>
<td>406</td>
<td>0.0006</td>
<td>0.04</td>
<td>0.16</td>
</tr>
<tr>
<td>Macroinvertebrate richness</td>
<td>TP (&lt; 1.6 mg/L)</td>
<td>230</td>
<td>&lt;0.0001</td>
<td>0.36</td>
<td>0.12</td>
</tr>
<tr>
<td>Sestonic chlorophyll-a</td>
<td>TN (&lt; 3.0 mg/L)</td>
<td>233</td>
<td>&lt;0.0001</td>
<td>0.11</td>
<td>0.07</td>
</tr>
<tr>
<td>Fish richness</td>
<td>TN (25.0 mg/L)</td>
<td>1091</td>
<td>&lt;0.0001</td>
<td>0.08</td>
<td>12.1</td>
</tr>
<tr>
<td>Macroinvertebrate richness</td>
<td>Turbidity (≥200 NTU)</td>
<td>163</td>
<td>&lt;0.0001</td>
<td>0.47</td>
<td>13.7</td>
</tr>
</tbody>
</table>

Table 15. Breakpoint values for several significant (p<0.05) linear, piecewise regression models based on different stressor (independent) and response (dependant) variables taken from the USEPA Region 7 RTAG stream database.

1 Includes only sites having at least one sensitive species.
2 One site (TP = 1.82 mg/L) was excluded from the analysis.
3 228 sites (TN >3 mg/L) were excluded from the analysis.
4 Four sites (TN >25 mg/L) were excluded from the analysis.
5 Four sites (NTU > 200 NTU) were excluded from the analysis.
in the water column (Figure 23). These data suggest threshold nutrient concentrations (levels at which we see a sharp increase in the probability that benthic chlorophyll will exceed 100 mg/M²) occur in the range of 20-80 μg/L TP and 200-800 μg/L TN (Figure 23). These plots allow visualization of the probability that benthic chlorophyll will exceed some level given a specific concentration of nutrients in the water column. In general, mean values of chlorophyll exceeding 100 mg/M² and maximum above 150 mg/M² are considered excessive (Dodds and Welch, 2000).
The Region 7 RTAG met on December 12, 2006, and agreed on benchmarks to protect uses and biotic integrity. We discussed several issues related to how suggested levels should be applied: 1) Should stream size matter in criteria?; 2) Should criteria be ecoregion-specific?; 3) What role should observed thresholds in nutrient versus biotic responses play?; 4) Should turbidity be considered in nutrient criteria?; and 5) How do the various methods (thresholds, reference streams, trisection, and USEPA 25%) compare?

The first issue was whether criteria should be set specific to stream or river size. In general, there was little relationship between watershed size and response variables (TN, TP, turbidity, benthic chlorophyll). The only exception was that sestonic chlorophyll was greatest in the largest watersheds. The consensus was that for recommending nutrient criteria levels, there was little to be gained from separating out larger rivers from others. Sample size for rivers was also relatively small for all nutrient variables because few states had active large river monitoring programs in place.

The second issue was the question of ecoregion-specific criteria. The main hurdle in this discussion was the fact that some ecoregions are highly impacted and have few reference sites. In addition, it was recognized that several ecoregions were poorly represented in Region 7, with only a small spatial extent of the ecoregion occurring within the region’s four-state area. Trisection and USEPA 25% methods could be greatly influenced by lack of good reference sites and a preponderance of highly nutrient-enriched sites. The question then became: Can we determine reference levels in the absence of pristine or native sites? Several literature approaches have allowed this, and the literature data were discussed. The ecological thresholds also are important in this determination. We could find no scientific justification for different responses to nutrients across ecoregions. This conclusion was buttressed at least in part by the literature (Dodds et al.,...
by human uses that the technological limitations or socioeconomic factors might make it impractical to lower nutrients to levels protective of uses or aquatic life. However, the RTAG did not consider this point because their specific charge was to determine the scientific basis behind levels for nutrient criteria, not the feasibility of regulating to reach some set level of nutrient criteria in terms of what were historical regions. Socioeconomic and implementation issues related to the adoption of nutrient criteria were issues that fell beyond the expertise and charge of the RTAG. These issues are better addressed by economists, engineers, stakeholders and decision-makers. We simply could not address technical limitations of nutrient criteria except for those that would be attainable given historic conditions.

In the end, the group could find little scientific justification for proposing criteria that varied by ecoregions within the USEPA Region 7 states and tribal lands. The point was made that reference conditions set a lower bound on nutrient criteria because you cannot hope to lower nutrient concentrations below those that were historically present before the advent of widespread anthropogenic influences (e.g., before widespread fertilization of cropland, numerous large confined animal feeding operations, and substantial populations releasing sewage effluent into waterways).

The third issue discussed was what the thresholds mean. The group decided that threshold levels of TN and TP above which chlorophyll no longer increases substantially provide an indication of regions where nutrient control may be effective. That is, with respect to benthic algal biomass, there will be limited biological response to different nutrient criteria if nutrients remain above threshold levels. Numerous thresholds also were identified where macroinvertebrate biodiversity decreased with increasing nutrients, but then these thresholds stabilized at low levels above some threshold nutrient concentration. Given the probability that this threshold level indicates a range above which most organisms have little exposure over evolutionary time, macroinvertebrate thresholds were viewed as an upper level for protecting aquatic life. Development of nutrient thresholds for fish variables were limited because few significant piecewise regressions were found deviations from chlorophyll/TN response only in subtropical ecoregions. While basic stressor/response relationship did not differ significantly between ecoregions, regional differences in stressor concentrations, both anthropogenic and natural, may require establishment of benchmark/criteria at some finer geographical scale (e.g. Rohm et al., 2002).

A question of attainability related to the ecoregion has come up repeatedly as part of the discussions of nutrient criteria. Some regions are already so heavily impacted by human uses that the technological limitations or socioeconomic factors might make it impractical to lower nutrients to levels protective of uses or aquatic life. However, the RTAG did not consider this point because their specific charge was to determine the scientific basis behind levels for nutrient criteria, not the feasibility of regulating to reach some set level of nutrient criteria in terms of what were historical regions. Socioeconomic and implementation issues related to the adoption of nutrient criteria were issues that fell beyond the expertise and charge of the RTAG. These issues are better addressed by economists, engineers, stakeholders and decision-makers. We simply could not address technical limitations of nutrient criteria except for those that would be attainable given historic conditions.

In the end, the group could find little scientific justification for proposing criteria that varied by ecoregions within the USEPA Region 7 states and tribal lands. The point was made that reference conditions set a lower bound on nutrient criteria because you cannot hope to lower nutrient concentrations below those that were historically present before the advent of widespread anthropogenic influences (e.g., before widespread fertilization of cropland, numerous large confined animal feeding operations, and substantial populations releasing sewage effluent into waterways).

The third issue discussed was what the thresholds mean. The group decided that threshold levels of TN and TP above which chlorophyll no longer increases substantially provide an indication of regions where nutrient control may be effective. That is, with respect to benthic algal biomass, there will be limited biological response to different nutrient criteria if nutrients remain above threshold levels. Numerous thresholds also were identified where macroinvertebrate biodiversity decreased with increasing nutrients, but then these thresholds stabilized at low levels above some threshold nutrient concentration. Given the probability that this threshold level indicates a range above which most organisms have little exposure over evolutionary time, macroinvertebrate thresholds were viewed as an upper level for protecting aquatic life. Development of nutrient thresholds for fish variables were limited because few significant piecewise regressions were

### Table 19: Summary of values relevant to setting benchmark values in streams and rivers.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Literature(^1,2,3) (ranges)</th>
<th>Nutrient Regions(^4) (range)</th>
<th>Reference Streams (median)</th>
<th>Trisection(^5) (median)</th>
<th>MEANS (all methods)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN (mg/L)</td>
<td>0.7 – 1.5(^5) 0.15 – 1.10(^5) 0.51 – 0.54(^3)</td>
<td>0.54 – 2.18</td>
<td>1.08</td>
<td>0.81</td>
<td>0.82</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>0.025 – 0.075(^5) 0.023 – 0.060(^2) 0.027 – 0.043(^3)</td>
<td>0.01 – 0.128</td>
<td>0.08</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Sestonic Chla (μg/L)</td>
<td>10 – 30(^1)</td>
<td>0.9 – 3.0</td>
<td>3.3</td>
<td>2.8</td>
<td>2</td>
</tr>
<tr>
<td>Benthic Chla (mg/M²)</td>
<td>20 – 70(^1)</td>
<td>NA</td>
<td>24.2</td>
<td>20.3</td>
<td>11.9</td>
</tr>
<tr>
<td>Turbidity (NTU)</td>
<td>NA</td>
<td>1.7 – 17.5</td>
<td>12</td>
<td>10.5</td>
<td>9.5</td>
</tr>
</tbody>
</table>

\(^1\)Dodds et al., 1998  
\(^2\)Dodds and Oakes 2004 and Smith et al., 2003  
\(^3\)Dodds et al., 2002  
\(^4\)From EPA 822-B-00-017, -18, -019, -020, EPA 822-B-01-013, -014, -016.  
\(^5\)Trisection values are for upper one-third streams in US EPA Region 7 having highest total richness for macroinvertebrates.

### Table 20: The final proposed nutrient benchmark values and response values most likely to protect use and integrity of waters for median values of TN, TP, and sestonic and benthic chlorophyll a for EPA Region 7.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Benchmark value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TN (mg/L)</td>
<td>0.9</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>0.075</td>
</tr>
<tr>
<td>Sestonic Chla (μg/L)</td>
<td>8</td>
</tr>
<tr>
<td>Benthic Chla (mg/M²)</td>
<td>40</td>
</tr>
</tbody>
</table>
identified. Those that were significant had low $R^2$ values, and threshold values were two to 12 times higher than breakpoints for macroinvertebrate variables (Table 15). However, in Ohio streams, Miltner and Rankin (1998) found a negative correlation between nutrients and their biotic index for fish. Increasing nutrient concentrations in low-order streams were associated with deleterious effects on the fish community, especially when nutrient levels exceeded background (total inorganic nitrogen and phosphorus > 610 and 60 μg/L, respectively). These numbers are similar to RTAG benchmark numbers for TN and TP. Based on data from 70 wadeable streams within watersheds exhibiting a large gradient of agricultural land-use practices, a TN threshold of 480 μg/L was identified by Maret et al. (2010), where eutrophic index scores based on aquatic plant metrics abruptly became less responsive to increasing TN concentrations.

The role of stream turbidity in altering the relationships between nutrient stressors and receptors could not be determined. Consistent relationships between turbidity and other factors related to nutrient criteria were poorly delineated in our dataset, except for the very strong positive relationship between TP and turbidity. Consequently, the group decided not to consider turbidity in its final recommendations for benchmark values, but observed that methods to control turbidity also control non-point sources of total phosphorus.

The final area of discussion was how the various methods compare (Table 19). The group felt that there was general concordance across methods and thus chose values for TN, TP, sestonic Chl$\alpha$, and benthic Chl$\alpha$ that were obtainable given baseline nutrient concentrations for the ecoregions. These values were consistent with the numbers determined from nearly all of the multiple lines of evidence used by the RTAG. Values were greater than the thresholds for biodiversity, consistent with the group’s understanding of how these thresholds should be used. The selected benchmark values (Table 20)
are remarkably similar to other proposed benchmark and criteria values listed in Table 19. The point was made with regard to these benchmarks that they apply only to streams and rivers. The values are higher than those chosen for lakes in Region 7, and managers may need to set more restrictive values if protection of downstream waterbodies is required.

Summary

As mentioned, the group urges the use of information on five weight-of-evidence factors (reference conditions, historical data and trends, models, RTAG expert review and consensus, and downstream effects), as well as information from the literature, establishing nutrient water quality criteria. These elements, as expressed in USEPA’s technical manuals (USEPA, 2000a; 2000b; 2001b) ideally should be incorporated in the criteria development process. The RTAG of the USEPA Region 7, States, and Tribes are the most knowledgeable parties for incorporating this information into comprehensive criteria development. In the absence of this effort, USEPA may be obliged to rely extensively on the benchmark values presented in this report for any necessary nutrient quality management decision-making. Thus, States are strongly encouraged to use this information as their basis for developing criteria that are more geographically specific and refined. With these benchmarks for decision-making, USEPA-state cooperation can be established to protect our rivers and streams.
In conclusion, the RTAG produced a series of violin plots comparing the relationships between proposed regional benchmark values to the values for all sites (e.g., streams), as well as for a priori reference sites within only those ecoregions having ≥5 sites, each with five or more samples (Figures 24-27). It appears that in about half to one-third of the ecoregions, stream populations had median or 25th quartile values that were at or below the TP and TN benchmark threshold levels (Figures 24 and 25). All but two of the seven ecoregions that had enough reference site data to plot had median TP values below the benchmark concentration of 75 μg/L. In all or nearly all ecoregions that had sufficient data to construct violin plots, these plots had median and 25th quartile levels well below the sestonic and benthic chlorophyll benchmark values of 8 and 4 μg/L, respectively (Figures 26 and 27). Only the stream populations (i.e., all streams) of the Central Great Plains (CGP), Central Irregular Plains (CIP), Flint Hills (FH) and Interior River Lowlands (IRL) ecoregions had median sestonic Chlα levels above the 8 μg/L benchmark concentration. These comparisons help illustrate the general applicability of these regional benchmarks and point out the need for more data to refine benchmark values in some areas of USEPA Region 7, especially those ecoregions that occur mostly outside of Region 7.
Literature Cited


