HALL & ASSOCIATES

Suite 203
1101 15th Street, N.W.
Washington, D.C. 20005-5004
Telephone: 202-463-1166
Web Site: http://www.hall-associates.com
Fax: 202-463-4207
Reply to E-mail:

jhall@hall-associates.com

August 21, 2008

Via Email and U.S. Mail
Administrator Stephen L. Johnson
Environmental Protection Agency
Ariel Rios Building
1200 Pennsylvania Avenue NW
Washington, DC 20460

RE: Request for Peer Review of New EPA Region III Approach to Developing Instream Standards for Nutrients

Dear Administrator Johnson:

On behalf of the Communities located in three Pennsylvania watersheds (Chester, Paxton and Indian Creeks) and in accordance with the OMB's document entitled *Final Information Quality Bulletin for Peer Review*, we are requesting that EPA initiate a formal independent peer review of the unprecedented scientific approach EPA Region III utilized to develop stringent phosphorus stream standards for three major Pennsylvania ecoregions. The new standards will result in billions of dollars in additional treatment costs in Pennsylvania. Moreover, the standards are physically unattainable by the MS4 communities and non-point sources. As compliance with these requirements is impossible to achieve, severe economic impacts are expected to occur if these objectives are enforced. Consequently, it is imperative that the technical validity for this new approach be independently assessed.

The Region's unprecedented approach for developing stream nutrient standards does not consider whether nutrient control will limit plant growth, contrary to all published EPA Section 304(a) criteria development guidance for nutrients. The new approach *assumes* phosphorus directly impairs sensitive invertebrate populations regardless of its impact on aquatic plants, a concept previously rejected in EPA's nutrient criteria guidance. Moreover, the new statistical method (conditional probability) employed to generate the instream standard neither (1) confirms nutrients are the cause of any impairment nor (2) demonstrates that the Region's new use impairment of concern (invertebrate diversity) will be remedied by nutrient control. In short, the new procedures are a radical departure from published criteria development methods and historical approaches used to generate

Section 303(c) and 304(a) criteria that have always been premised on a clear scientific demonstration of causation and need.

The new technical procedures employed by the Region have never been documented or peer reviewed by EPA to be a sufficient or scientifically defensible basis to generate a necessary and protective Section 304 (a) numeric stream criterion. Nonetheless, the Office of Water informally determined that the new approach was a proper basis for developing Section 304(a) criteria and for imposing expansive point and non-point control measures for phosphorus reduction. (Attachment, Exhibit 8) Once our group realized the Office of Water was intent on supporting the Region's unprecedented approach, we contacted two internationally recognized experts, Dr. Dominic DiToro and Dr. Stephen Chapra, who EPA has often relied upon regarding nutrient issues, to ask for their opinion on the new approach. Their opinions were offered without request for compensation and were made solely due to their professional interest in ensuring scientifically defensible and cost effective approaches are used in nutrient regulation. Both experts clearly and very adamantly concluded the new approach and resulting numeric water quality standards (1) are not scientifically defensible, (2) would not likely restore stream impairments, and (3) would misdirect local resources. (Attachment, Exhibits 9 and 10) In particular, Dr. Chapra concluded that "This is such a scientifically indefensible representation of the connection between nutrients and ecosystem health that I believe that its adoption would represent a grave mistake.... I am much more concerned that its adoption would ultimately be ineffective. That is, it could lead to costly controls that would not protect our precious stream ecosystems." Both scientists stated that either a Science Advisory Board or National Academy of Sciences independent peer review of the new approach should occur given the new procedure's radical departure from accepted methods for criteria derivation and nutrient impact assessment.

Federal Policy and Guidance Mandate Peer Review

As you are aware, EPA has historically conducted an independent peer review of new scientific approaches before utilizing such approaches in the water quality criteria development process. (See, e.g., Science Advisory Board review of EPA's Approach to Emerging Contaminants and EPA's 2006 Peer Review Handbook). As stated in EPA's Handbook, "The principle underlying the Peer Review Policy is that all influential scientific and technical work products used in decision making will be peer reviewed." (Handbook @ 30) The purpose of the independent peer review is, at a minimum, to ensure EPA is basing its regulatory program requirements on scientifically-defensible and well-documented evidence linking the environmental concern to a workable regulatory solution. Virtually every criterion identified in EPA's Handbook (@ 30-32), to determine if a peer review is necessary, is met in this case: the technical work product uses a new scientific approach establishing a precedent that departs significantly from prior agency approaches and has widespread regulatory implications and cost impacts.

Moreover, OMB has adopted mandatory review procedures that are applicable in this instance given the controversy, high cost, and precedent setting nature of the Region III and Office of Water actions. (See, OMB Bulletin – *Final Information Quality Bulletin for Peer Review,* January 14, 2005). EPA plainly intends for the new procedures to be used on a nationwide basis for deriving stream nutrient standards, as acknowledged in the Office of Water memorandum. We understand that states have been requested to use this new approach in developing nutrient standards; therefore, such procedures constitute "highly influential scientific assessments" subject to the strictest peer review requirements. This new standards derivation procedure should have been peer reviewed *before* it was applied to Pennsylvania communities. Nonetheless, it is still subject to the OMB peer review requirement and that action should be promptly undertaken. (See OMB Bulletin, 70 Fed. Reg. 2673-2674, January 14, 2005)

The nationwide application of these new methodologies will result in stream nutrient standards that are beyond the limits of technology and are unattainable in most instances. The cost of compliance will easily exceed several hundred billion dollars. Given the nationwide importance of having scientifically-defensible procedures for generating stream nutrient standards, we respectfully request that you direct the Office of Water to submit the new approach for independent peer review at the National Academy of Sciences. Moreover, as it is highly probable that the new, more restrictive phosphorus reduction requirements mandated by EPA Region III's recent TMDL actions will not result in any meaningful ecological improvements, we ask that you stay application of those TMDLs in the NPDES process until the required peer review is completed.

The attached report, "Justification for Peer Review of EPA's New Methodology for Nutrient Criteria Development", provides additional background information on the history of the Region's TMDL development in Pennsylvania. Thank you for your consideration of this request.

Sincerely,

John C. Hall

Attachments

cc: Senator Arlen Specter
Senator Robert P. Casey
Benjamin Grumbles, USEPA
Pennsylvania Periphyton Coalition
Pennsylvania Municipal Authorities Association
Pennsylvania Water Environment Association
Home Builders Association of Metropolitan Harrisonburg

JUSTIFICATION FOR PEER REVIEW OF
EPA'S NEW METHODOLOGY FOR NUTRIENT CRITERIA DEVELOPMENT

Background on EPA Region III Nutrient Criteria Development in Pennsylvania

Since 2005, EPA Region III has encountered repeated difficulties in developing a scientifically defensible approach to determine the effects of nutrients on plant growth in streams. These difficulties, as discussed below, led the Region to abandon accepted methodologies for determining whether nutrients were causing stream use impairment (i.e., excessive plant growth). The new approach assumes that nutrients have a direct effect on invertebrate populations, as if this constituent was a toxic substance. A statistical method known as conditional probability was used to assess field data wherein it was assumed that nutrients were the direct cause of any monitored changes in invertebrate populations. Using the results of this analysis in a "weight-of-evidence" approach (that included three other methods that were divorced from any showing that nutrient levels caused changes in invertebrate populations), the Region selected a "growing season" total phosphorus level and asserted it would ensure the impaired invertebrate populations in the stream would be restored. The available site-specific data on invertebrate populations and nutrient levels were not used to assess whether or not the new approach produced a rational, scientifically defensible result.

The following evaluation reviews applicable regulatory requirements for criteria derivation and then details the fundamental changes the Region and EPA Headquarters made to those procedures in their effort to justify nutrient reduction for stream discharges. These fundamental changes were undertaken using new scientific methodologies that have never undergone any independent scientific peer review to determine their sufficiency for generating appropriate and protective stream nutrient standards.

Statutory and Regulatory Requirements for Criteria Derivation

By statute, criteria must be based on the latest available science and set at the level necessary to protect aquatic life and human health uses. CWA Section 304(a). To achieve this requirement it is essential that criteria possess two attributes: (1) the criteria must be based on data that confirm the pollutant is causing use impairment at ambient concentrations, and (2) the level at which the numeric criteria is set is both sufficient and necessary to protect stream uses. Thus, criteria are, in general, set at the threshold level where the pollutant exposure is demonstrated not to pose a significant threat to aquatic life. (Section 304(a); 40 CFR 131.2 131.3 (b), (c))

Since 1985, EPA has had a well defined procedure for developing water quality criteria when it published the document entitled "*Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses,* " USEPA 1985 (hereafter "*Guidelines*"). The *Guidelines* establish a number of very specific scientific thresholds that must be met to establish criteria that meet Section 304(a) mandates, as follows:

- Water quality criteria must ensure use protection "with a small probability of considerable overprotection or under-protection." (*Guidelines* @ 5)

- It is not enough that the criterion is the best estimate given the available data. Criteria should be derived "only if adequate appropriate data are available to provide reasonable confidence that it is a good estimate." (*Guidelines* @ 5)

- Criteria must be based upon studies showing a clear dose/response relationship to determine effect concentration. Data from confounded studies (i.e., results that are influenced by factors other than the pollutant of concern) should not be used. (*Guidelines* @ 15, 16, 21)

- All decisions should be based on a thorough knowledge of aquatic toxicology and criteria decisions must be altered if there is a substantial probability of over or under protection of aquatic organisms and their uses. (*Guidelines* @ 18)

- Based on "all available laboratory and field information," it must be determined that proposed criteria are "consistent with sound scientific evidence." If not, another criterion should be derived. (The concluding recommendation of the *Guidelines* @ 57)

The basic scientific premise underlying all published EPA nutrient criteria development documents is that nutrient control is intended to reduce excessive plant growth. Consistent with the *Guidelines'* requirements for a clear demonstration of causation, the various EPA nutrient criteria documents for lake and stream environments all clearly specify that dose/response demonstrations are required to set scientifically defensible nutrient standards. Nutrient levels must be documented to "cause" specific changes in plant growth (typically measured as chlorophyll 'a') and other physical variables directly affected by excessive plant growth (Secchi depth, DO, transparency, etc.). See, *Nutrient Criteria Technical Guidance Manual – Rivers and Streams*, USEPA July 2000 (hereafter "*Rivers and Streams Document*"). The *Rivers and Streams Document* is clear that a nutrient criterion must be based on a demonstration that nutrients are causing excessive plant growth (eutrophication). For example, Chapter 1 identifies various ecosystem impacts related to excessive plant growth, and specifies that nutrient criteria are based on the relationship between nutrient levels and plant growth as measured by chlorophyll 'a' ("Nutrient criteria development should relate nutrient concentrations in streams, algal biomass and changes in ecological condition (e.g., nuisance algae accrual rate and deoxygenation). … Initial criteria should be verified and calibrated by comparing criteria in the system of study to nutrients, chl *a* and turbidity values in water bodies of known condition to ensure that the system of interest operates as expected.") *Rivers and Streams Document* @ 13) "Predictive relationships between nutrients and periphyton (or phytoplankton) biomass are *required* to identify the critical or threshold concentrations that produce nuisance algal biomass." Id @ 76, emphasis supplied.

The various EPA nutrient criteria documents all acknowledge that nutrients may cause ecosystem impacts to upper level organisms (invertebrates, fishes), but *never* directly:

> "However, fish and macroinvertebrates do not directly respond to nutrients, and therefore may not be as sensitive to changes in nutrient concentrations as algal assemblages. It is recommended that relations between biotic integrity of algal assemblages and nutrients be defined and then related to biotic integrity of macroinvertebrate and fish assemblages in a stepwise, mechanistic fashion." *Rivers and Streams Document* @ 85.

EPA's published guidance indicates that invertebrate populations may be affected only when plant growth rises to a level where extensive/excessive plant growth causes those ecosystem changes. These changes are not documented to occur directly due to nutrients as this parameter is not a toxicant and does not have a direct impact on sensitive organisms. See, *Development and Adoption of Nutrient Criteria into Water Quality Standards*, USEPA 2001 @ 14, response 4. This fact was also well documented by EPA's field studies under the whole effluent toxicity program.

The *Guidelines* is quite clear that a simple "weight-of-evidence" approach is not a sufficient basis for setting a criterion. Furthermore, the *Guidelines* provide that a simple regression approach between two variables (one a field response) would not suffice as a demonstration that the input variable caused the effect measured in the field. Nowhere does this document indicate that a conditional probability approach may be used for derivation of a numeric standard. This is not unexpected since that statistical method cannot provide a demonstration that regulating a pollutant at a given level provides any assurance that use protection will or will not be achieved at that pollutant level. At best the method indicates the likelihood (i.e., the probability) of encountering the condition being evaluated for a given pollutant concentration used in the regression. Finally, both the *Guidelines* and the *Rivers and Streams Document* are replete with statements underscoring the need to understand the toxicology of the substance. To set a numeric standard, one must determine that the pollutant of concern is the direct cause of the adverse effect being measured. Simple regressions and conditional probability analyses provide no such confirmation.

Unless there is clear, well documented evidence on the level of water quality necessary and sufficient to protect aquatic life uses, criteria should not be established. Moreover, the *Guidelines* and the *Rivers and Streams Document* are clear that site-specific information should be considered if it shows that the suggested standard is misplaced. The new EPA approach expressly ignores such information. In each case where the new standards were applied in Pennsylvania, it was acknowledged that habitat degradation (sedimentation/channelization) was the root cause of any documented changes in invertebrate populations. Site-specific regressions were provided to demonstrate that, in fact, there was no relationship between nutrient levels and invertebrate populations in the various streams where such data were available. (Discussed below in greater detail). Contrary to EPA's own recommendations in the *Guidelines* and the *Rivers and Streams Document*, the Region simply ignored those data and analyses, claiming the new procedures provided sufficient confirmation that nutrients were the cause of stream impairments.

Published Literature Confirmed EPA's Simplified Approach to Stream Nutrient Regulation Was Misplaced

As EPA is well aware, the issue of how to develop scientifically defensible stream nutrient standards, in particular, has been a very controversial subject due to the technical challenges inherent in attempting to develop a uniform approach to such waters. It is now well established in the literature that plants inhabiting streams (periphyton/ macrophytes) do not respond as algal species inhabiting lakes (phytoplankton). For example, early on EPA often relied on research produced by Dodds, who sought to develop some type of simplified relationship between periphyton and nutrient levels. (See, *Rivers and Streams Document* @ 77; *Protocol for Development of Nutrient TMDLs, First Edition* – November 1999, USEPA 1999 @ 4-6). However, in Dodds' more recent publications he has concluded that periphyton growth has the capability of reaching very high levels even where very low nutrient levels are present. [1] Thus, it is now apparent that the simplified approaches do not work for controlling periphyton growth in streams as EPA originally had contemplated.

As noted in EPA's criteria documents, in many situations, nutrient levels do not control plant growth but other physical factors do. For periphyton, in particular, this seems to be the case. Likewise, EPA now recognizes that many macrophytes may obtain nutrients from the soil matrix, rendering control of water column nutrient levels a meaningless exercise. (See, *Rivers and Streams Document* @ 73)

The observations of Dr. Dodds were becoming apparent to many others around the country. On July 17, 2007 the Association of State and Interstate Water Pollution Control Administrators (ASIWPCA) sent a letter to the Office of Water informing EPA that states were unable to demonstrate the necessary relationships between nutrient levels and EPA's recommended instream response parameters (e.g., plant growth, turbidity, DO, etc.). (Exhibit 1) That organization raised critical concerns that continuing on the path chosen by EPA would invariably lead to wasted expenditures which state and local governments can ill afford:

> *These problems can only lead to miscues in impairment identification and mis-direction of scarce management and implementation resources…. Because no two water bodies are the same, site-specific evaluations and most probably, site specific criteria are required that reflect their uniqueness and protect their natural trophic tendencies. This will be a costly endeavor but less financially*

[1] Walter K. Dodds. 2006. "Eutrophication and trophic state in rivers and streams." Limnol. Oceanogr. 51(1, part 2) p 671 – 680. "[A]ttached algae might be able to attain impressive biomass in nutrient-poor water because periphyton can use the small amounts of nutrients that continuously flow by.". @ 677

costly than attempting to meet water quality criteria that are unattainable and less environmentally costly than losing water resources because criteria are too liberal.

Thus, ASIWPCA called on EPA Headquarters to reconsider the efficacy of ecoregion or state wide numeric nutrient standards in light of well documented problems in demonstrating that nutrient control would be effective in regulating excessive plant growth in streams. A focus on site-specific conditions was identified as the only way to ensure proper and effective programs for stream restoration. [2]

EPA Region III Mishaps with Nutrient Regulation

EPA Region III itself encountered these same difficulties identified by ASIWPCA in attempting to develop necessary and protective nutrient objectives for five watersheds across the Commonwealth of Pennsylvania. [3] After several years' effort, EPA withdrew the proposed restrictive phosphorus TMDL and instream standards it had developed in late 2006 for Wissahickon Creek. This TMDL sought to limit periphyton growth via point source controls. The site specific periphyton data for that stream, however, clearly documented two facts: (1) upstream of the wastewater plants where TP levels were quite low, periphyton growth was as robust as downstream where TP levels were quite high, and (2) periphyton growth was closely correlated to tree canopy, confirming that light, not nutrients, was the limiting factor for plant growth. [4] EPA coauthored a peer reviewed journal article (with its technical consultant Tetra Tech) that concluded nutrient reduction would not be effective in controlling plant growth because low TP levels could support robust periphyton growth. [5] ("However, it is worth mentioning that while periphyton activities was one of the major causes of the DO violations, it was finally determined to be infeasible to control the periphyton through reducing nutrient loads from point sources.")

Region III Employs an Unprecedented Approach to Nutrient Criteria Development and TMDL Decision Making

Due to the repeated problems encountered in attempting to relate nutrient levels to periphyton growth in streams, the Region employed a new, technically-unprecedented approach to develop stream nutrient standards. The new approach ignored whether nutrient levels affected plant growth and did not even attempt to demonstrate whether nutrients were actually causing any site-specific aquatic life (invertebrate) impairment in

[2] Establishing a clear relationship between nutrient loading and plant growth is not generally problematic in lake environments, where such relationships have been well documented for decades.

[3] EPA detailed the problems it encountered in attempting to develop a periphyton/nutrient relationship in a May 2, 2007 affidavit filed by Thomas Henry, USEPA, in the matter of American Littoral Society v. EPA.

[4] Alan Everett. February 19, 2002. Periphyton Standing Crop and Diatom Assemblages in the Wissahickon Watershed. Montgomery and Philadelphia Counties. @ 12

[5] Zou et al,."Integrated Hydrodynamic and Water Quality Modeling System to Support Nutrient Total Maximum Daily Load for Wissahickon Creek, Pennsylvania," Journal of Environmental Engineering, April 2006.

the streams at issue. Rather, the Region acquired phosphorus and invertebrate population data for three "ecoregions" and simply had the data plotted, with total phosphorus or total nitrogen as the independent variable and various invertebrate metrics as the dependent variable. All measured changes in invertebrate populations were assumed to be a direct result of the nutrient concentration exposure. [6] EPA Region III, using the same contractor (Tetra Tech – see footnote 5), developed three "regional" (Eastern Piedmont, Central, and Allegheny Plateau) total phosphorus standards as part of five promulgated nutrient TMDLs in Pennsylvania. The numeric nutrient standards ranged from 25 to 40 ug/l as "growing season" (April to October) averages. (See, *Development of Nutrient Endpoints for the Northern Piedmont Ecoregion of Pennsylvania*, Tetra Tech, November 20, 2007; hereafter "*Tetra Tech Report*" and *Development of Nutrient Endpoints for Allegheny Plateau and Ridge and Valley Ecoregions of Pennsylvania: TMDL Application*, Tetra Tech, June 24, 2008.)

In setting these standards, Region III interpreted Pennsylvania's narrative nutrient standard using an approach that deviated from the prior nutrient impact assessment guidance established by Pennsylvania and EPA that required a demonstrated linkage between increased nutrient levels and excessive plant growth. [7] In place of a stream model such as QUAL2 or WASP5, a statistical procedure known as "conditional probability" was used in a "weight-of-evidence" analysis to generate the instream numeric water quality standard. The conditional probability procedure simply plotted datasets from a selected ecoregion that measured instream TP levels and assessed the type of invertebrate populations present. Several different invertebrate indices were chosen as the endpoint of concern (e.g., total taxa, # EPT taxa, % clingers, etc.) for comparison with total phosphorus and total nitrogen levels. A sample of that analysis is presented in Exhibit 2. It is worthy to note that the "weight-of-evidence" analysis evaluated multiple macroinvertebrate metrics that were deemed an important measure of aquatic health and discarded those that did not show sensitivity to TP rather than factor those results into the analysis (See, *Tetra Tech Report* @18).

---

[6] This technical assumption was at odds with prior EPA scientific conclusions regarding the manner in which nutrients affect the environment. EPA's nutrient criteria guidance specifically states that macroinvertebrates ***do not*** respond directly to nutrients. *Nutrient Criteria Technical Guidance Manual – Rivers and Streams*, EPA-822-B-00-0002, Ch. 6, pg. 85 (July 2000).

[7] Pennsylvania DEP's guidance on assessing nutrient impairment to free-flowing streams requires documentation of excessive plant growth caused by increased levels of phosphorus or nitrogen and attendant violations of the dissolved oxygen standard. *Implementation Guidance for Section 95.9 Phosphorus Discharges to Free Flowing Streams*, Department of Environmental Protection Bureau of Watershed Conservation, (October 27, 1997) at pg. 7. ("For purposes of this guidance, a nutrient-related problem is defined as a documented use impairment due to nuisance algal or rooted aquatic plant growth conditions with attendant violations of dissolved oxygen standards."). The designation of a stream as nutrient impaired requires a similar demonstration. *Commonwealth of Pennsylvania Assessment and Listing Methodology*, 2004/2006: *Cause Definitions – Nutrients* ("The presence of excessive quantities of Phosphorus and/or Nitrogen that under the proper conditions may result in dense algal or macrophyte growth and wide fluctuations in Dissolved Oxygen levels.") Regarding the TMDL actions at issue, EPA Region III filed a sworn affidavit with the Court stating that Pennsylvania law required EPA to demonstrate how nutrients impacted plant growth in order to establish a lawful basis to regulate nutrients. (See, n. 3)

Contrary to the *Guidelines* and the *Rivers and Streams Document*, no attempt was made to show that TP was the actual parameter causing the change in invertebrate populations or indices, or to show confounding factors did not influence the invertebrate metrics in EPA's database. Tetra Tech's analysis of the impaired water bodies acknowledged that factors other than phosphorus were causing the changes in invertebrate populations being measured, but EPA had informed Tetra Tech that the waters were nutrient impaired so Tetra Tech's analysis ignored a causal assessment.[8] Tetra Tech also acknowledged that the nutrient levels chosen to protect invertebrate populations *would not* limit plant growth since the selected target was well above published limiting nutrient levels. [9] The method employed was a mere correlation that could not and did not show that nutrients were the cause of the any changes in invertebrate indices. [10]

No attempt was made to demonstrate that the chosen indices were set at a level necessary to protect the stream uses. [11] The acceptable invertebrate levels were simply selected as the midpoint of the rating scale used by the state of Maryland. Based on the selected metric, 50% of the sites used by EPA to evaluate the nutrient standard would be considered impaired or non-attainment sites. Many of these impaired stations have TP levels well below the target TP value established by EPA, confirming that factors other than phosphorus are significantly influencing this database. The most this methodology could indicate is that there was a very weak relationship ($R^2 < 0.1$) [12] between total phosphorus and sensitive invertebrate populations. The R^2 for the data was quite low (~ 0.1 – that is less than 10% of the data response is explained by TP levels). To discern the "protective" standard from this analysis, the contractor looked at the "change point" in the graph. This is the location in the graph where it becomes more probable that the selected instream metric will not be attained. Even at the selected "instream standard" there was a 50% probability that the target invertebrate levels would not be achieved. [13]

---

[8] Response Document for Nutrient and Sediment TMDLs in Pennsylvania for Southampton Creek, Indian Creek, Chester Creek, Paxton Creek, and Sawmill Run, June 30, 2008 @ 14. "Again, this effort was not undertaken to "show" that TP is the cause of impairment … Tt (sic "Tetra Tech") was not asked to determine the cause of impairment; we were given a cause and asked to determine a protective value."

[9] *Development of Nutrient Endpoints for the Northern Piedmont Ecoregion of Pennsylvania: TMDL Application*. November 20, 2007. at 15-16. "Not surprisingly, a strong algal biomass-nutrient relationship was not present in our examination of the data sets … Surprisingly, the highest algal biomass occurred as sites where the TP concentrations were relatively low, 14 – 35 ug/L. It is possible that algal growth has been saturated even at this low level."

[10] *Id*. at 11. "Correlation analyses identified significant relationships between biological response and nutrient variables. However, correlation may or may not indicate the real relationship. Numerous relationships were examined; only a subset of which was correlated. There were also results that were considered potentially important but showed weaker relationships (Appendix A)."

[11] In an August 4, 2008 Pennsylvania Right-to-Know Law response, PA DEP stated that it had no documentation or other information showing that the invertebrate target endpoints were necessary or appropriate levels for assessing stream impairments.

[12] *Id* at 18.

[13] Hall and Associates duplicated the Tetra Tech calculations and evaluated the change points associated with varying EPT Taxa conditional probabilities (i.e., <1 – <12 EPT Taxa). The calculations resulted in virtually identical TP concentrations regardless of the conditional EPT Taxa level. To most, this shows that the instream target is not a function of the phosphorus concentration. Incredibly, EPA asserted that this

Nonetheless, both Tetra Tech and EPA concluded that the selected standard would ensure that uses were fully protected and invertebrate impairments restored by meeting this value.

Because EPA's use of the methodology was unprecedented and a radical shift in the established analytical framework, numerous questions were raised in various public meetings held by EPA to present the TMDLs. In these public meetings and subsequent FOIA responses, EPA and its contractor (Tetra Tech) acknowledged that:

- The approach *did not prove* phosphorus was causing the invertebrate response.

- Meeting the chosen numeric nutrient standard would *not ensure* restoration of the target invertebrate population.

- There was no demonstration that the chosen instream metrics were *necessary* to provide use protection; the selected metrics were just the median values from the scoring criteria.

- The approach was not based on any *site-specific information* demonstrating a relationship between elevated TP levels and invertebrate populations.

- The approach did not consider the available site-specific data which, in general, confirmed that factors other than phosphorus were the root cause of the changing invertebrate levels (e.g., habitat alteration).

Regarding the issue of site-specific stream impairment data, EPA applied the new numeric standards to one watershed that was never identified as nutrient impaired on any TMDL list (Chester Creek). The periphyton data provided by EPA as part of the final TMDL confirmed that plant growth in Chester Creek was rather minimal and well below the level EPA thought could cause adverse impacts (i.e., > 150-200 mg/m^2 as a growing season average). Various biological assessments had determined that habitat impairment caused reduced invertebrate populations in one segment of the stream (Goose Creek). This assessment was not considered in applying EPA's new nutrient criteria as the solution to the problem. EPA also ignored the fairly extensive and only site-specific invertebrate data for that watershed, which it had included in the TMDL document. Those data confirmed that phosphorus levels in allegedly impaired segments of the watershed were *unrelated* to invertebrate populations. (Exhibit 3) [14] In fact the *highest* phosphorus levels in the "impaired" segment were associated with the *best* invertebrate

---

complete lack of dose/response confirmed that nutrients are a dramatic stressor. (Supra Note 8 at 17-18) The EPA response is contrary to accepted scientific principles.

[14] EPA's response to comments on the TMDL ignored this analysis finding that the Tetra Tech report was a sufficient basis to conclude that phosphorus was in fact impairing Chester Creek. This position violates the *Guidelines* which requires reconsideration of an approach if the field data confirms it is misplaced (see, *Guidelines* @ 57).

population readings that surpassed the impairment threshold used in the *Tetra Tech Report*. Moreover, the chosen TP standard was violated uniformly in both upstream and downstream waters that DEP had determined fully attained uses. Thus, the new recommended standard applied to Chester Creek does not differentiate between waters with acceptable invertebrate levels and waters with allegedly unacceptable levels.

EPA also applied the new invertebrate impacts-based standard to Paxton Creek. The lower section of this stream that was the focus of the nutrient TMDL is *concrete lined.* Consequently, habitat was well documented as very poor and invertebrate levels were, as expected, quite reduced. Invertebrate levels were robust in other unimpaired segments of the creek, despite having TP levels far in excess of the instream standard that Region III claimed was necessary to protect uses. Analyses of these site-specific data show that the invertebrate populations are much more correlated to habitat and very poorly correlated to phosphorus levels (Exhibits 4 and 5, respectively). As with Chester Creek, the chosen standard (25 ug/l TP growing season average) could not distinguish between impaired and unimpaired segments of the creek. (Exhibit 6) The good correlation between invertebrate population and habitat score confirmed the overwhelming importance of this stressor. However, EPA ignored the ramifications of these data, finding that the new instream TP standard was a valid indicator of impairments due to nutrients.

EPA Headquarters Review Supports The Region's New Approach and Confirms that Nationwide Implementation Of a New Nutrient Criteria Development Procedure is Being Promoted

Given the apparent inconsistency with the *Guidelines* and prior EPA Nutrient Criteria development documents, the inconsistency with the site-specific information, and the complete lack of documentation showing that TP was actually the pollutant causing the changes in invertebrate levels, the group of affected Pennsylvania communities approached EPA Headquarters in April 2008 to request an independent review. Various letters were sent to Benjamin Grumbles and the Office of General Counsel. Initially, the staff's informal response was that the new approach was not consistent with Section 304(a) criteria development requirements and that conditional probability could not be used as the basis to derive a numeric water quality standard. After two months, however, it became apparent that the Office of Water was intent on supporting the new Regional approach to nutrient criteria development using conditional probability and assessing nutrient impacts as if nutrients were toxicants.

To understand the rationale behind this decision, a Freedom of Information Act (FOIA) request was sent to EPA Headquarters to obtain any available background documentation supporting EPA's position. EPA's June 19, 2008 FOIA response confirmed the following. Apparently, after receiving the letter from ASIWPCA, rather than address the difficult technical issues raised (i.e., plant growth in streams is not well connected to nutrient levels), the Office of Water decided to use a new nutrient standard approach that ignored whether or how plant growth was affected by nutrients. A conditional probability approach presented in a paper entitled "*Development of empirical, geographically specific water quality criteria: a conditional probability analysis* *approach*" Paul and McDonald (2005), Journal of American Water Resources Association 41:1211-1223 was now the recommended basis for deriving nutrient standards based on invertebrate impacts. (Exhibit 7) [15] In August, 2007 EPA began a nationwide series of "workshops" under the Agency's "N Steps program" to launch its new criteria derivation approach and convince states that this radical new approach was scientifically defensible. Thus, via this series of internal presentations, EPA completely abandoned the published national nutrient criteria approaches specified in the Agency's guidance (e.g., the *Rivers and Streams Document*.) [16] No public notice or peer review of this new approach was given prior to this radical change in nutrient criteria derivation procedures.

It should be noted that Dr. Paul's (EPA ORD) presentation materials entitled "Conditional Probability Analysis: A Statistical Analysis Tool," as well as the original paper co-authored by Dr. Paul (Exhibit 7), contained the following cautions regarding use of conditional probability based upon field data to identify an appropriate instream standard:

> **Disadvantages**: Other stressors confound the association
>
> **Other Points to Remember**
>
> Conditional probability is just a statistical tool that can be used to extract very specific information from a data set. Before applying CPA ("change point analysis"), *it is imperative* that *extensive* laboratory data analysis (EDA) be conducted. EDA is a form of detective work, primarily using graphical depictions of various renditions of the data. (emphasis supplied)

Dr. Paul's examples only suggested applying this procedure in the area of the data where there was great certainty that the stressor was highly correlated to the impairment. EPA left out these cautions in its subsequent September, 2007 Regional Nutrient Criteria Development Workshops.

A June 11, 2008 memorandum by William Sweitlik, Chief, USEPA Ecological and Health Processes Branch, Office of Science and Technology, confirmed that the Regional Office had simply followed EPA Headquarters new advice and should be commended for implementing the procedures in the TMDLs. (Exhibit 8). [17] The memorandum stated the following:

[15] In October 2003 the EPA Science Advisory Board considered but did not agree that the conditional probability approach was an appropriate methodology for setting suspended and embedded sediment standards. The procedure was noted as a useful tool, though not sufficient to derive a numeric standard. It was noted that any use of this methodology had to be based on a documented strong stressor response.

[16] The only reference to prior EPA approaches that required a clear demonstration of how nutrients impacted plant growth and then ecosystem indicators was a slide entitled: *Nutrients…ughh.*

[17] William Swietlik to Robert Koroncai. June 11, 2008. Development of Nutrient Endpoints for TMDLs in Pennsylvania.

> It is our conclusion that the approach used in the document (sic –*Tetra Tech Report*) … is a scientifically defensible approach and is consistent with EPA guidance for deriving nutrient criteria. The approach used in the document is an example of the multiple-lines-of-evidence (or weight-of-evidence) approach. … In October 2007 EPA HQ provided training to the Region II and III States on the weight-of-evidence methodology and how it can be applied to developing numeric nutrient values. It is good to see the Region benefited from our training and you are now employing this approach.

Thus, it is apparent that EPA Headquarters has launched a new method for nutrient criteria derivation and that it expects it to be used on a nationwide basis. Discussions with other communities across the country verified that, in fact, state agencies are being requested to use this new approach for standards development.

Among the documents alleged to support the new methods was a prior Science Advisory Board review of the conditional probability method. The development documents specify that the metric for which a criteria is developed using conditional probability <u>must be a strong stressor</u> (i.e., the aquatic community condition is clearly related to the stressor for higher values of the stressor). The background documents cautioned that there must be a clear, scientifically established causal relationship between the pollutant at issue and the endpoint selected for review. [18] The Tetra Tech conditional probability analysis expressly excluded any demonstration that nutrients were, in fact, the cause of any documented change in invertebrate populations. Tetra Tech informed the public that it was directed by EPA to *assume* that the changes in invertebrate populations were caused by phosphorus concentration. [19]

Contrary to the admonitions of Dr. Paul (USEPA ORD) and the Science Advisory Board, the site-specific information for the streams at issue were either unavailable or not considered to determine if TP was a strong stressor or if other factors were confounding the finding that nutrients were the culprit. The site-specific data for Chester Creek and Paxton Creek clearly confirmed TP level was <u>not</u> a strong stressor, if at all. [20] Thus, it is apparent that the Office of Water's

---

[18] The prior consideration of the conditional probability method was for stream sediment impacts. It is well documented that impactedness and sedimentation may severely degrade invertebrate habitats in streams. There is no such scientific demonstration regarding nutrient concentrations. To the contrary, it is well documented that nutrient do not directly impact invertebrates or their habitat.

[19] See, USEPA Region III response to comments on Nutrient TMDLs for Chester, Paxton and Indian Creek. ("Again, this effort was not undertaken to "show" that TP was the cause of impairment. … Tt (sic "Tetra Tech") was not asked to determine the cause of impairment; we were given a cause and asked to determine a protective value.")

[20] The Swietlik Memorandum supported using multiple-lines-of-evidence and cited EPA's support of this approach in guidance documents, including "*Nutrient Criteria Technical Guidance Manual: Rivers and Streams*." In citing this guidance, Headquarters stated that the weight-of-evidence approach combines several approaches including: 1) reference reaches, 2) predictive relationships, and 3) published threshold values. However, the predictive relationships identified by Region III were confounded as standard

support for the Region's approach did not even follow its own guidance on whether such a method may be considered for identifying an appropriate instream standard, when applied to nutrients.

Moreover, while the *Rivers and Streams Document* does discuss considering "multiple-lines-of-evidence" as a way to *strengthen* a scientifically defensible finding, nowhere does that document or the *Guidelines* suggest that a criterion is scientifically defensible simply because it uses "multiple-lines-of-evidence or weight-of-evidence." Finally, nowhere has any published, peer reviewed nutrient criteria development approach stated that it is acceptable to (1) assume impacts are caused by a pollutant, (2) ignore whether plant growth will be affected by nutrient regulation, (3) assume nutrients directly impact invertebrates without documented laboratory studies confirming that fact, or (4) ignore site-specific information that shows nutrient regulation is not necessary or will be ineffective. These new EPA assumptions are radical departures from published, scientifically defensible procedures EPA has used for decades under the Section 304(a) criteria development and 303(d) TMDL programs.

Request for Independent Peer Review

The new approach to developing numeric nutrient standards for streams is scientifically unprecedented and a radical departure from published EPA criteria development methods. If this standards derivation methodology remains unchanged, dischargers throughout Pennsylvania (and eventually the country) will be required to install extremely advanced phosphorus treatment at exorbitant costs with little likelihood of producing demonstrable environmental benefits. While our coalition understands that environmental expenditures will be necessary to ensure that our lakes and rivers meet their designated uses, they are very wary of using their limited resources in an unnecessary fashion or a manner that will not produce the desired results.

EPA Headquarters Office of Water has apparently promoted and now approved the radical new nutrient criteria derivation approach. This new approach has never undergone the peer review or technical evaluation process required of all EPA criteria development changes. For the reasons detailed below, we request that EPA promptly conduct an independent peer review of the new EPA nutrient standard setting approach using either EPA's Science Advisory Board or the National Academy of Sciences.

First, it is apparent that the new approach is contrary to a series of "bedrock" scientific principles relied upon by the Office of Water for decades, including:

---

development tools because "other stressors" exert a greater impact (see *Tetra Tech Report* at 15- 21). As for macroinvertebrates, Region III reported that three of the six metrics considered where either not sensitive to nutrient enrichment or more sensitive to other stressors. Of the remaining three, the regression coefficients were extremely poor. Finally, reference reaches upstream of the municipal facilities confirmed that low invertebrate populations were caused by other stressors, <u>not</u> nutrients. Thus, the primary assumption required for using conditional probability (i.e., the aquatic community condition is clearly related to the stressor for higher values of the stressor) was not met.

- Numeric criteria must be based on documented dose/response relationships between the pollutant and a use impairment (versus assuming the pollutant is causing the problem and ignoring data to the contrary)
- Numeric standards must be set at the level found both necessary and sufficient to protect uses (versus setting the standard where the probability of impacts is decreased even if the stressor response is extremely weak)
- Nutrients are not directly toxic to invertebrates but affect plant growth (versus ignoring the degree of plant growth occurring and assuming that nutrients directly impact invertebrate populations)
- Confounded data may not be used to develop a numeric standard (versus assuming all measured field responses are due to a pollutant, even where the data show this is not true), and
- Site-specific data, when available, must be considered in determining whether a numeric standard is necessary and will achieve its intended level of protection (versus ignoring the site-specific data and assuming that the generalized conditional probability analysis is accurate in all cases)

EPA may not legitimately abandon well established scientific principles and requirements and alter its published criteria development approaches by simply hosting an "ad hoc" series of workshops that recommend that accepted approaches be changed. The public has an absolute right under the Clean Water Act to participate in such critical decision making of nationwide importance. EPA's current approach is contrary to basic principles of administrative law and lacks the transparency that is required of all major regulatory decisions. See, CWA Section 101(e) ("Public participation in the development, revision and enforcement of any regulation, standard, effluent limitation, plan or program established by the Administrator or any State under this Act *shall be provided for*, encouraged and assisted by the Administrator and the States.") There is hardly a more important program than that used to establish the basic water quality criteria for protection of the Nation's waters. Under Section 304(a)(3), …"Such criteria and information *and revisions thereof, shall be issued to the states and shall be published in the Federal Register and otherwise made available to the public*." The Clean Water Act plainly does not contemplate that major changes to criteria development procedures are to be clandestinely launched via internal EPA workshops and announced to the public as a *fait accompli* under the TMDL program.

Second, federal peer review procedures require that new, innovative or controversial scientific procedures used to establish regulatory program requirements must first undergo peer review before they are used in a regulatory context. EPA has long had a peer review process applicable to changes in criteria derivation methods. (See USEPA Peer Review Policy, 1993) This is a typical situation that would have to undergo federal peer review under EPA's own guidance. In fact, the criteria derivation method employed in this case, and to be employed nationwide, failed to receive Science Advisory Board approval when last considered by that peer review panel. Moreover, on December 16, 2004, the Office of Management and Budget (OMB) issued a final bulletin to all agencies establishing that influential scientific information shall be peer reviewed before it is disseminated by the Federal government. (70 Fed. Reg. 2664, January 14, 2005) EPA updated its own peer review policy to accommodate the OMB requirements (EPA/100/B-06/002, May 2006). Although agencies have discretion to choose the specific type of peer review to employ, the duty to conduct a peer review is not discretionary. *Id.* at 2675. In determining the extent of the peer review necessary, the OMB bulletin stated that "[m]ore rigorous peer review is necessary for information that is based on ***novel methods*** or presents complex challenges for interpretation. Furthermore, the need for rigorous peer review is greater when the information contains ***precedent-setting*** methods or models, presents conclusions that are likely to ***change prevailing practices***, or is likely to affect ***policy decisions*** that have a significant impact." *Id.* at 2668. (emphasis added). There is no serious question that EPA's attempt to use a new scientific approach to nutrient criteria derivation, at odds with its published scientific approach, meets every component of the OMB Bulletin justifying peer review.

Third and finally, the opinions of two internationally renowned experts in the field of nutrient control (who have voluntarily reviewed Region III's approach at our request) state that the approach used and endpoints derived will not ensure designated uses are met, and that using conditional probability to establish the endpoints was not scientifically defensible. (See, letters of Drs. Di Toro and Chapra, Exhibits 9 and 10) [21] These scientists clearly state that EPA's approach is not based on accepted scientific principles and should be peer reviewed by the Science Advisory Board or the National Academy of Sciences, as has occurred in similar cases where new, scientific approaches are being employed in the regulatory program.

As outlined above, the approach used by Region III to set nutrient endpoints in the recently released Pennsylvania TMDLs warrants an independent peer review. Specifically, the new approach is precedent-setting, uses novel methods, and will change prevailing administrative practices. Beyond that, international experts on the issue believe that the approach taken is misguided. Finally, if not modified, the potential cost impact to the Pennsylvania dischargers will be on the scale of billions of dollars, and the nationwide potential to misdirect resources is virtually certain to occur. Misdirection of resources will result in unabated environmental impairments and excess energy usage unrelated to environmental need. The new approach would, in all likelihood, cause more harm than good.

As such we respectfully request that EPA initiate a Science Advisory Board or National Academy of Sciences independent peer review on this new procedure for deriving

[21] Dr. Dominic Di Toro and Dr. Stephen Chapra, both internationally recognized experts on environmental pollution matters, state that EPA's approach was (1) not scientifically defensible, (2) did not demonstrate TP was causing any impairment (3) could easily regulate the wrong pollutant (TP instead of sedimentation) and, in any event, (4) did not ensure that the chosen standards would protect the stream uses. Both support that this new procedure should undergo an independent peer review *before* it is used in a regulatory context. The Region had Dr. Di Toro's letter prior to the completion of the TMDL and simply ignored it. However, despite repeated requests, Region III has proffered no credible peer reviewed studies showing that total phosphorus acts like a toxicant and directly impacts sensitive invertebrate populations.

numeric nutrient criteria for streams. Pending that review, it is also respectfully requested that EPA stay further application of this methodology as well as implementation of the TMDLs that were derived using this unauthorized method because EPA failed to follow the statutory requirements of CWA § 304(a) prior to relying on this unorthodox criteria derivation methodology.

EXHIBIT 1

ASIWPCA

Association of State and Interstate Water Pollution Control Administrators

1221 CONNECTICUT AVENUE N. W, 2ND FLOOR, • WASHINGTON, DC 20036 • TEL: 202-756-0600 • FAX: 202-756-0605 • WWW.ASIWPCA.ORG

Ben Grumbles
Assistant Administrator for Water
US Environmental Protection Agency
1201 Pennsylvania Avenue, N.W.
Room 3219
Washington, DC 20469

July 18, 2007

SUBJECT: Nutrient Pollution and Numeric Water Quality Standards Memo

Dear Mr. Grumbles,

The Association of State and Interstate Water Pollution Control Administrators (ASIWPCA) and its Member States received your memorandum of May 25, 2007, which urges accelerated promulgation of numeric water quality criteria for nutrients and the establishment of additional controls for nutrient pollution. Member States agree that nutrient controls are a critical and necessary component of comprehensive water quality management. In addition to limiting or eliminating discharges of priority pollutants and toxics in treated wastewater effluent, management of water quality is shifting to include controls on less traditional parameters such as sediments and excess nutrients.

However, these parameters also exist as part of a balanced natural aquatic system that is often dynamic and does not exhibit "threshold" effects that are amenable to generic numeric criteria setting. Clearly we are entering a new era of water quality management in which the traditional approach of criteria, permits and enforcement, needs to be reevaluated to ensure its pertinence to Water Quality Standard attainment. This is especially true for nutrients and related response variables that exhibit a wide range of conditions representative of a diversity that must be maintained. The uniformity of eutrophic and productivity conditions that numeric criteria would promote defies both common sense and basic principles of ecological succession, which define homeostasis in the natural world.

The May 25 memo also raises several issues that we believe should be addressed cooperatively by States and EPA. In summary:

- EPA should continue to refine and enhance the scientific basis for numeric nitrogen and phosphorus criteria. Many States are not finding a scientific link between cause and effect that is needed to support numerical standards.
- EPA must not undercut EPA approved State nutrient criteria development plans including their agreed upon milestones and commitments.
- A nutrient criterion requires appropriate implementation procedures. This is a costly endeavor, but well worth it.
- Direct impacts on permittees in the NPDES program need to be considered.

- EPA should support quantitative economic assessment of numerical nutrient controls and management that go beyond local watersheds and State boundaries to include upwind and downstream relationships.
- Given EPA's high level of concern about nutrients, the Agency should develop categorical standards for POTWs and have consistent realistic national effluent limits, rather than relying on a State-by-State battle with poorly supported numerical water quality standards.
- The table in the memo on the status of activities in the States should be updated and clarified.

ASIWPCA strongly encourages EPA continue to refine and enhance the scientific basis for numeric nitrogen and phosphorus criteria, including reevaluating the potential difficulties with statistically-derived generic criteria that may be over or under protective. During their considerable developmental processes, many States are failing to find a strong linkage between the EPA recommended cause variables (N and P) and response variables of chlorophyll-a and transparency, but are finding wide variations in parameters that seem unrelated to professional assessments of "trophic health" status. In many cases, a relationship cannot be demonstrated between causal variables N and P, and factors such as turbidity, light limitation, canopy cover, substrate, aquatic community structure, bioavailability, reservoir sequestration, micronutrient limitations and other "response" variables. These problems can only lead to mis-cues in impairment identification and mis-direction of scarce management and implementation resources.

We emphasize that most States have a "mutually agreed upon" nutrient criteria development plan approved by EPA that includes commitments and specific milestones for the development and promulgation of both narrative and numeric criteria. The May memo seems to override those agreements which States would strongly oppose. While States are actively working to meet the mutually agreed upon milestones of these plans, many are improving water quality through the implementation of their new or existing narrative criteria, often using the TMDL process to meet related criteria, e.g., for dissolved oxygen or aquatic life use support. Nutrient control programs are not waiting on a number and are in fact well underway in every State.

Implementation is a key component of effective water quality management. By itself, the singular act of promulgating a criterion does nothing to improve water quality. A criterion requires appropriate implementation procedures to support its application. Consideration of mixing zones, reasonable potential tests, averaging periods, assessment methodologies, dose-response relationships, and other related factors are all absolutely critical if criteria are to be effectively set and used for developing effluent limits, TMDL targets, non-point source control practices and use impairment decisions for Integrated Reports. Because no two waterbodies are the same, site-specific evaluations and, most probably, site-specific criteria are required that reflect their uniqueness and protect their natural trophic tendencies. This will be a costly endeavor, but less financially costly than attempting to meet water quality criteria that are unattainable, and less environmentally costly than losing water resources because criteria are too liberal.

States are acutely aware that numeric nutrient criteria, whether generic or site-specific, will ultimately have a direct impact on permittees in the regulatory NPDES program, with correspondingly significant monetary impacts. Therefore, in addition to ensuring that the criteria are appropriate to the water body, the States need more capability to quantify the costs and benefits of nutrient controls as they progress toward implementing nutrient control plans. It must be recognized that the practice of promulgating water quality criteria without correspondingly considering implementation is outdated. States also recognize that nutrient pollution in many watersheds is largely or exclusively attributed to non-point sources and may involve both aqueous and aerial transport across jurisdictional boundaries. Control costs in these programs are much more difficult to quantify, and management efficiencies are very uncertain, in both the near and long term. All stakeholders want to know not only WHAT the number is, but HOW, WHERE, and WHEN it will be used in all phases of water quality management, and that the number is representative of a desired outcome for that water body. The EPA criteria documents referenced in the memo fall far short of meeting those needs, instead falling back on simplistic statistical categorizations that ill-consider the local character and capacities of individual water bodies and their typologies.

States encourage EPA to support a quantitative economic assessment of numerical nutrient controls and management that extends beyond the local watershed and State boundary to include upwind and downstream relationships. Nutrient impacts, and management benefits, in larger watersheds are often far reaching. For example, the economic costs of nutrient controls in the upper Midwest will be significant with little to no ecological improvement within that specific watershed, but they are critical to controlling eutrophication in estuaries such as the Chesapeake Bay or hypoxia in the Gulf of Mexico. Without a larger scale perspective it will be very difficult for States to justify the expenditure of significant financial resources when there will be little or no water quality improvement in their State. Therefore, States will require more resources to coordinate, facilitate and resolve interstate Water Quality Standards issues related to nutrient controls and reductions. This comes at a time when States are receiving less 106 funds for their core water program.

More time and attention should be paid to establishing nationwide nitrogen and phosphorus "effluent guidelines" or similar guidance. Wastewater treatment technology is making significant advances with state of the art phosphorus removal approaching 0.02 mg/L in contrast with present expectations (i.e. 1 mg/L). The effort needs to be pursued with great care since reliability is an issue and costs can be enormous for required filtration media to meet ultra-low levels. Energy and carbon issues are also a concern. In an era of "Water Quality Based Effluent Limits", it seems that technological controls should be a national USEPA priority to protect near shore marine environments, the Great Lakes, bays, estuaries and the Gulf of Mexico.

The table in the May 25, 2007 memo outlining the current status of numerical criteria adoption appears to be out-dated and in need of revision. Several States have pointed out that while they are actively engaged in developing nutrient criteria it was not reflected in the table. Since any table of this sort will be "old" within months of being compiled, it should be qualified so the reader understands that and carefully worded so that States get credit for making progress, not criticism. The table should reference States progress toward meeting objectives, goals and milestones in their approved nutrient criteria development plan.

Thank you for the opportunity to share our thoughts and ideas on this critically important issue. The Association looks forward to discussing this issue further with the Agency. Please do not hesitate to contact me at 405-530-8800.

Sincerely,

Derek Smithee

Derek Smithee, Co-Chair
ASIWPCA Monitoring, Assessment and Standards Task Force

Cc: Ephraim King, USEPA
Denise Keehner, USEPA
Grace Robiou, USEPA
Amy Newman, USEPA
Mary Smith, USEPA
ASIWPCA Membership

EXHIBIT 2

Conditional Probability

N-STEPS Objectives

Provide regions, states, and tribes with support related to nutrient criteria development

Provide access to expert assistance with issues related to nutrient criteria development and implementation

Improve communication nationwide.

What Is It?

A conditional probability is the probability of an event occurring when some other event has occurred and conditional probability analysis is an approach that allows a user to estimate the likelihood of exceeding some response threshold for a given nutrient concentration. This approach is based on the assumption that as nutrient concentrations increase, the likelihood of an impact on some negative response increases. This approach has great promise in identifying appropriate nutrient criteria to other measures where criteria thresholds have already been developed - for example, dissolved oxygen, pH, or biocriteria.

Example Question: What is the probability of exceeding our dissolved oxygen criterion over a range of different nutrient concentrations and where is the threshold?

How is it Applied to Nutrient Criteria Development?

Nutrient criteria development involves three main processes: identifying relationships between biological responses and nutrient stressors, examining these relationships, and establishing nutrient and/or biological thresholds or criteria.

Conditional probability analysis (CPA) is statistical tool to evaluate the relative risk of impairment of biological attributes along a nutrient gradient. It calculates the probability of exceeding a given threshold (e.g., biocriterion) given a set nutrient concentration; it can be used, therefore, to identify which nutrient conditions have a significant probability of being associated with adverse biological conditions. By combining CPA with change-point analysis, one can identify the nutrient threshold indicating a high risk of exceeding some other criterion.

How Does It Work?

Conditional probability calculates the probability of an event occurring (e.g., DO<y) when it is known that some other event has occurred (e.g., TP>x). The nutrient concentrations are treated as discrete random variables and probability functions are calculated. Functionally, a two-step procedure is used. One identifies the subset of samples were nutrients exceed some threshold and from those sites, one determines the samples which have exceeded the response threshold. This is the subset of samples for which nutrients exceed some value (x), which are also impacted for the response variable (y). Calculating these values iteratively over the range of nutrient concentrations generated an empirical conditional probability curve. Confidence intervals can be generated for this curve using resampling techniques, like bootstrapping (Paul and McDonald 2005). Thresholds along this curve are identified using variations of change-point analysis (See other fact sheet on change-point analysis).

Data Requirements

Independently collected numeric data in the form of paired observations are required. These are preferably continuous data, although discrete numeric variables (e.g., taxa richness) could also work. The greater the range of environmental conditions encompassed the better. One way to assure a large range is to use a gradient design and select sites along as large a gradient as possible.

What Should You Look For & Report?

Examine the conditional probability plots. They should show a clear trend in response with changes in nutrient concentration. Steep relationships and clear changes in the response make threshold identification easier. Confidence intervals can be used to help identify thresholds (see Paul and McDonald 2005). Confidence intervals for the change-points can also be reported (see fact sheet on Change-Point Analysis), depending on the approach used.

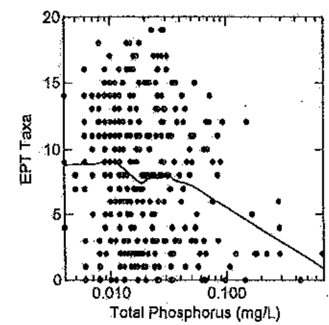
Pros	Cons
• Effective way to identify nutrient thresholds • Nice approach for linking criteria together	• Requires substantial data, probabilistic data is best • Lack of significance does not mean lack of association
• Quantitative measure of thresholds with error estimate	• Requires knowledge of an appropriate threshold for response variable • Lack of significance does not mean lack of association

Alternatives

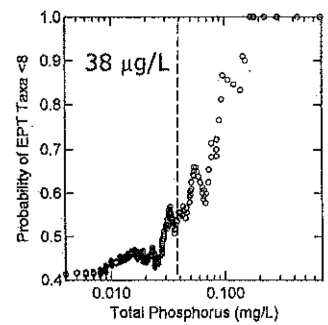
- Change-point analysis alone

Citations

Paul, J.F. and M.E. McDonald. 2005. Development of empirical, geographically specific water quality criteria: a conditional probability analysis approach. Journal of the American Water Resources Association 41:1211-1223



Scatterplot with Loess Line



Conditional Probability Plot

For more information contact
Steve Potts
US EPA
202-566-1121
Potts.Steve@epa.gov

EXHIBIT 3

Chester Creek TMDL
Summary of Macroinvertebrate Metrics and Nutrient Concentrations in East Branch Chester Creek (1996/1997)

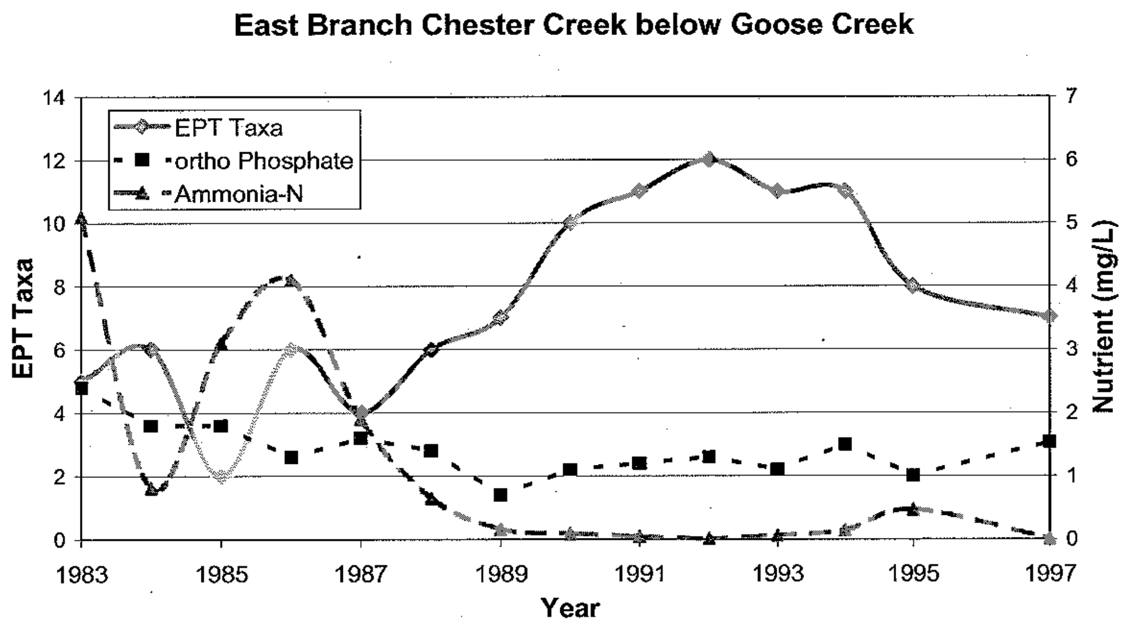
Station	Year	Total Taxa [1]	Total EPT Taxa [1]	HBI [1]	TN [2] (mg/L)	TP [2] (mg/L)
1476790	1997	17	9	4.2	4.3	0.018
1476830	1996	29	10	5.6	2.4	0.020
1476835	1996/1997	30/26	13/12	4.7/4.6	3.0	0.205

Definitions: EPT – Ephemeroptera, Plecoptera, and Trichoptera; HBI – Hilsenhoff Biotic Index; TN – total nitrogen; TP – total phosphorus

[1] Macroinvertebrate metrics from USGS Water-Resources Investigations Report 02-4242 "Assessment of Stream Conditions and Trends in Biological and Water-Chemistry Data from Selected Streams in Chester County, Pennsylvania, 1981-97" by Andrew G. Reif @ 60-62.

[2] Nutrient averages from Table 3-18, Nutrient Total maximum Daily Loads for the Chester Creek Watershed, Pennsylvania – Draft Report, February 2008.

Temporal Changes in EPT Taxa, Ammonia-nitrogen, and Ortho-phosphate at USGS Station No. 01476848 (East Branch Chester Creek below Goose Creek near West Chester)



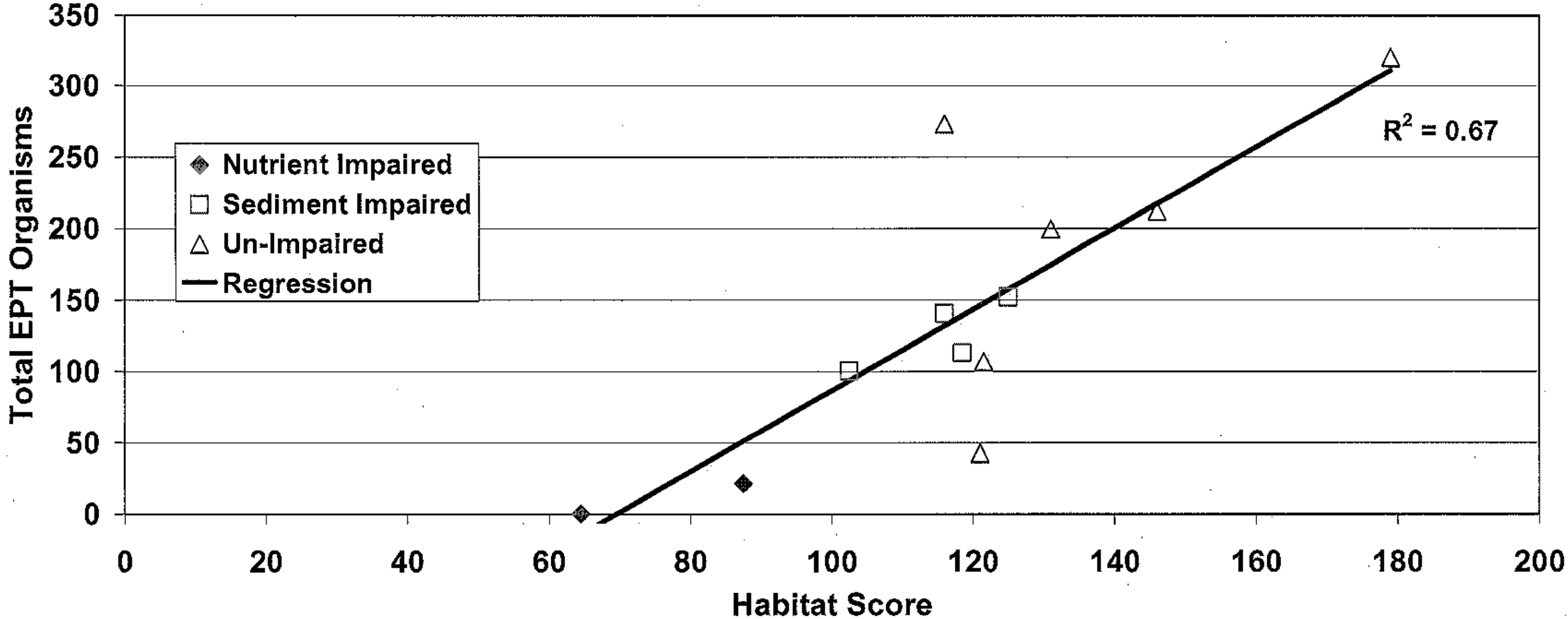
Macroinvertebrate metrics from USGS Water-Resources Investigations Report 02-4242 "Assessment of Stream Conditions and Trends in Biological and Water-Chemistry Data from Selected Streams in Chester County, Pennsylvania, 1981-97" by Andrew G. Reif @ 62.

Analytical data for ortho-Phosphate and Ammonia-nitrogen from USGS Open-File Report 99-216 "Physical, Chemical, and Biological Data for Selected Streams in Chester County, Pennsylvania, 1981-94 by Andrew G. Reif @ 127-128 and USGS Open-File Report 00-238 "Physical, Chemical, and Biological Data for Selected Streams in Chester County, Pennsylvania, 1995-97 by Andrew G. Reif @ 11.

EXHIBIT 4

Exhibit 4

Paxton Creek - 2006-2007 SRBC Data



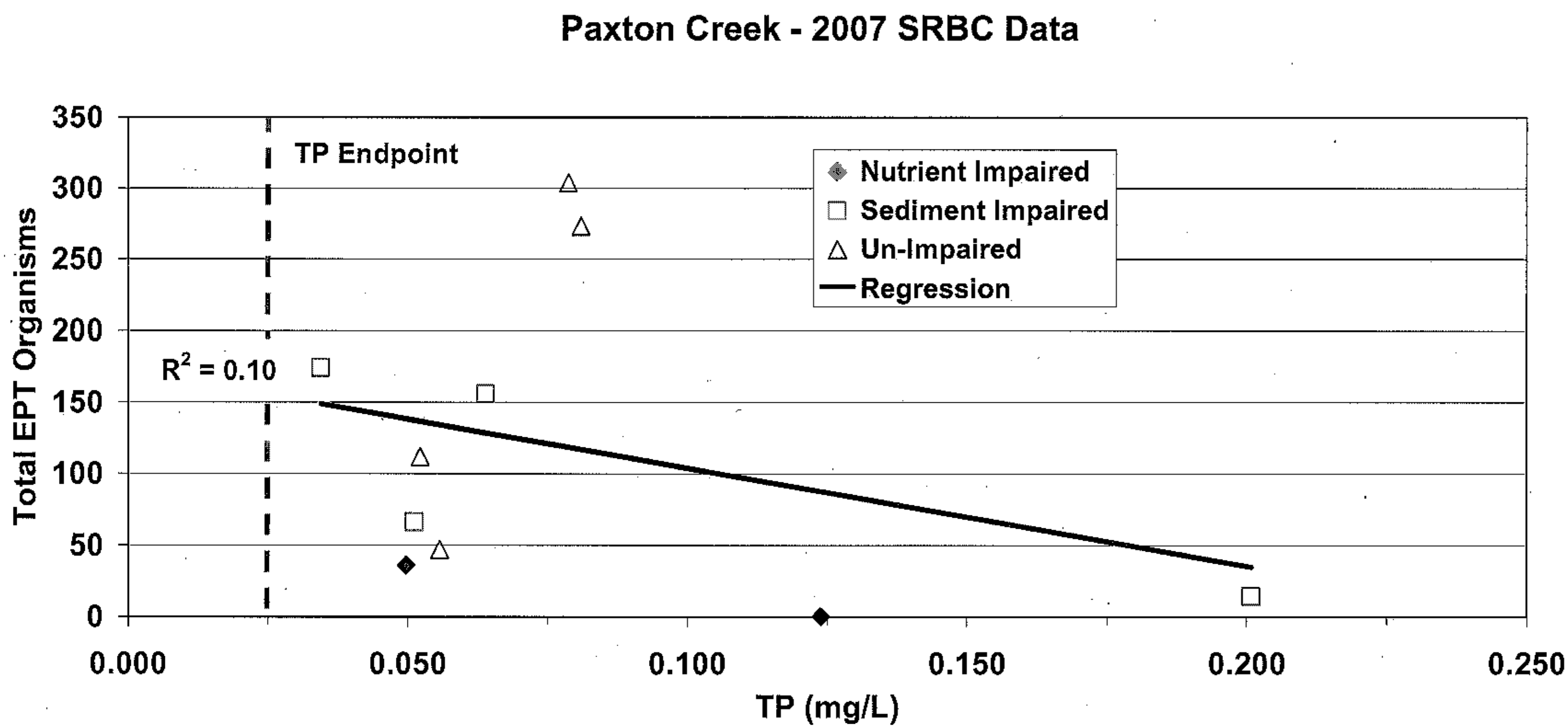
EPT – Ephemeroptera (Mayflies), Plecoptera (Stoneflies), and Trichoptera (Caddisflies)

Habitat scores from Nutrient and Sediment Total Maximum Daily Load in Paxton Creek Watershed, Pennsylvania (June 30, 2008) Table A-4 @ A – 23.

Macroinvertebrate metrics data (total EPT taxa) from Nutrient and Sediment Total Maximum Daily Load in Paxton Creek Watershed, Pennsylvania (June 30, 2008) Table A-5 @ A-25 and Table A-6 @ A-26.

EXHIBIT 5

Exhibit 5



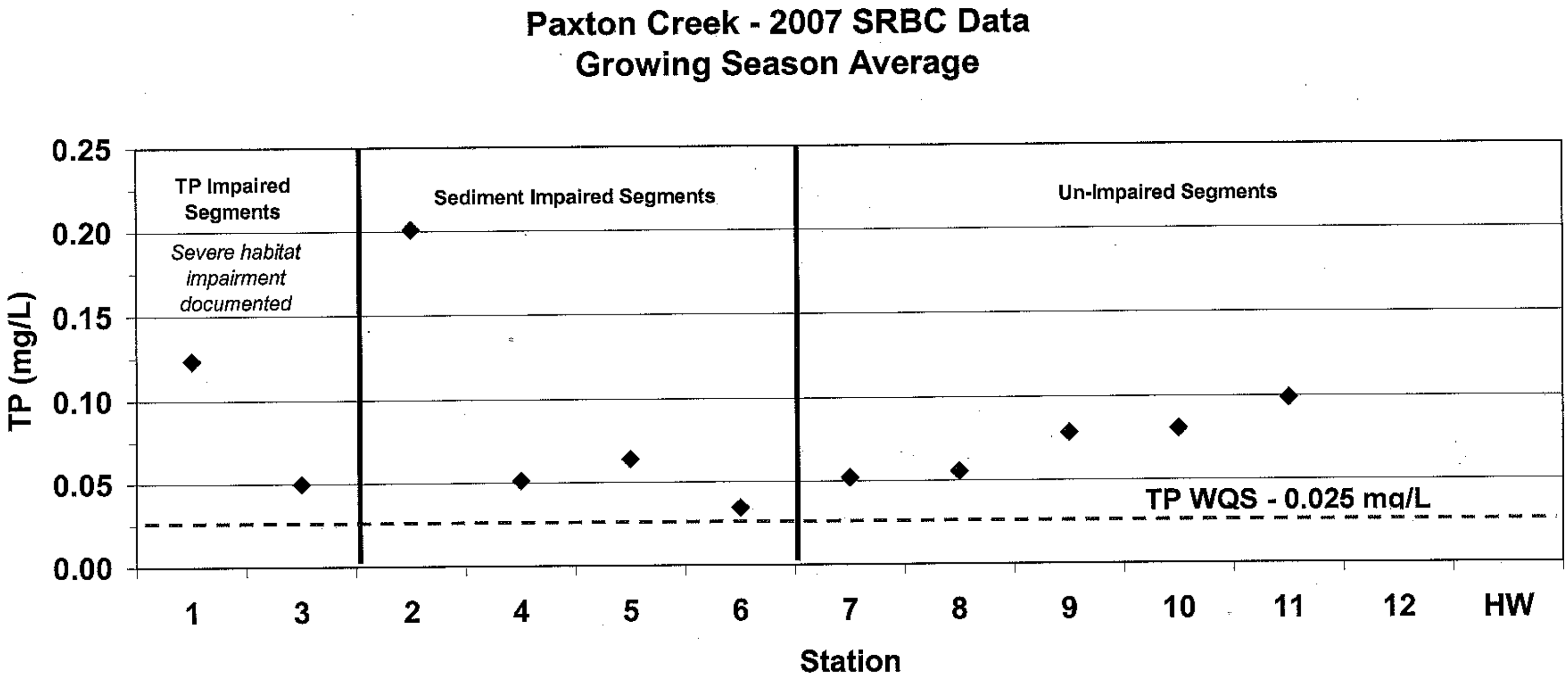
EPT – Ephemeroptera (Mayflies), Plecoptera (Stoneflies), and Trichoptera (Caddisflies)
TP – total phosphorus

Nutrient concentration data obtained from Response Document for Nutrient and Sediment TMDLs in Pennsylvania for Southampton Creek, Indian Creek, Chester Creek, Paxton Creek and Sawmill Run. June 30, 2008. @ 89 - 91

Macroinvertebrate metrics data (total EPT taxa) from Nutrient and Sediment Total Maximum Daily Load in Paxton Creek Watershed, Pennsylvania (June 30, 2008) Table A-6 @ A-26.

EXHIBIT 6

Exhibit 6



TP – total phosphorus
WQS – water quality standard

Nutrient concentration data obtained from Response Document for Nutrient and Sediment TMDLs in Pennsylvania for Southampton Creek, Indian Creek, Chester Creek, Paxton Creek and Sawmill Run. June 30, 2008. @ 89 - 91

EXHIBIT 7

JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION
OCTOBER AMERICAN WATER RESOURCES ASSOCIATION 2005

DEVELOPMENT OF EMPIRICAL, GEOGRAPHICALLY SPECIFIC WATER QUALITY CRITERIA: A CONDITIONAL PROBABILITY ANALYSIS APPROACH[1]

John F. Paul and Michael E. McDonald[2]

ABSTRACT: The need for scientifically defensible water quality standards for nonpoint source pollution control continues to be a pressing environmental issue. The probability of impact at differing levels of nonpoint source pollution was determined using the biological response of instream organisms empirically obtained from a statistical survey. A conditional probability analysis was used to calculate a biological threshold of impact as a function of the likelihood of exceeding a given value of pollution metric for a specified geographic area. Uncertainty and natural variability were inherently incorporated into the analysis through the use of data from a probabilistic survey. Data from wadable streams in the mid-Atlantic area of the U.S. were used to demonstrate the approach. Benthic macroinvertebrate community index values (EPT taxa richness) were used to identify impacted stream communities. Percent fines in substrate (silt/clay fraction, < 0.06 mm) were used as a surrogate indicator for sedimentation. Thresholds of impact due to sedimentation were identified by three different techniques, and were in the range of 12 to 15 percent fines. These values were consistent with existing literature from laboratory and field studies on the impact of sediments on aquatic life in freshwater streams. All results were different from values determined from current regulatory guidance. Finally, it was illustrated how these thresholds could be used to develop criterion for protection of aquatic life in streams.
(KEY TERMS: sediment; wadable streams; benthic community condition; statistical analysis; aquatic ecosystems; standards.)

Paul, John F. and Michael E. McDonald, 2005. Development of Empirical, Geographically Specific Water Quality Criteria: A Conditional Probability Analysis Approach. Journal of the American Water Resources Association (JAWRA) 41(5):1211-1223.

INTRODUCTION

A range of procedures are being used around the world for establishing criteria for the protection of water quality (Jimenez *et al.*, 1999; Yin *et al.*, 2003; Borja *et al.*, 2004; Kamizoulis and Saliba 2004, Kay *et al.*, 2004). In the United States, the U.S. Environmental Protection Agency (USEPA) is responsible for implementing the Clean Water Act (CWA) (Russo, 2002), which is the major national act for protecting water quality. The USEPA implements some aspects of the CWA by providing guidance for the control of pollutants through development of Water Quality Standards (WQS). These WQS serve as the foundation for pollution control and are a fundamental component of water quality management. They define the goals for a waterbody by designating its uses, setting criteria to protect those uses, and protecting water quality through antidegradation provisions. The criteria are developed for the protection of aquatic life as well as for human health.

Water quality criteria (WQC) for individual chemical pollutants (such as heavy metals and synthetic organic compounds) have been developed as national criteria (e.g., USEPA, 1994). These national criteria have been developed through laboratory bioassays, where exposure to a single pollutant can be maintained under controlled conditions (Hohreiter and Rigg, 2001; Rausina *et al.*, 2002; Fisher and Burton 2003). As progress has been made in controlling these individual pollutants, a shift has occurred toward control of nonpoint source pollution (e.g., runoff, nutrients, and sedimentation). Consistent with this is the increased use of biological indicators to assess the condition of the environment (Niemi and McDonald, 2004). Recent developments suggest that adopting national criteria may not be sufficiently protective of the biota in various subregions (Perry and Vanderklein, 1996; USEPA, 2000a), thus leading to a greater reliance on field generated data. The recent development of WQC for nutrients is an example of

[1]Paper No. 04095 of the *Journal of the American Water Resources Association* (JAWRA) (Copyright © 2005). **Discussions are open until April 1, 2006.**

[2]Respectively, Research Environmental Scientist and Director, Environmental Monitoring and Assessment Program, U.S. Environmental Protection Agency, Mail Drop 343-06, Research Triangle Park, North Carolina 27711 (E-Mail/Paul: Paul.john@epa.gov).

geographically specific criteria developed from field data (USEPA, 2000a).

Sedimentation in streams is an example of a non-point source pollution problem (Spooner *et al.,* 1991). Excessive sediment is a major cause of impairment in waterbodies across the country (USEPA, 2002). Development of water quality criteria for suspended and bedded sediments for the protection of aquatic life provides a challenge since the traditional approach using laboratory bioassays (dose-response studies) may not be applicable (Perry and Vanderklein, 1996).

A focus on the response of aquatic communities to sedimentation emphasizes protection of these communities from adverse effects of sedimentation and is consistent with the use of biological criteria. Establishing a criterion for sedimentation allows the source of excess sediments to be addressed for regulation or remediation of the problem. However, identification of the source of the sediments is not necessary for the development of criteria and is not discussed further.

In this paper, it is shown how a conditional probability analysis can be used with empirical, probabilistic monitoring data for aquatic resources to establish thresholds of impact for a stressor for a specified geographic area. Scientifically defensible thresholds are a necessary first step in establishing protective criteria by environmental managers. The approach is demonstrated by applying it to wadable streams in the mid-Atlantic region of the United States, and establishing thresholds for impact of sedimentation on the streams in this region. Finally, an illustration is presented on how these thresholds could be used to develop a criterion for protection of aquatic life in these streams from sedimentation.

DATA SOURCES

The field data used in this paper are available through the USEPA Environmental Monitoring and Assessment Program (EMAP) web site (USEPA, 2004). These data were collected from the mid-Atlantic region streams in 1993 and 1994 and include

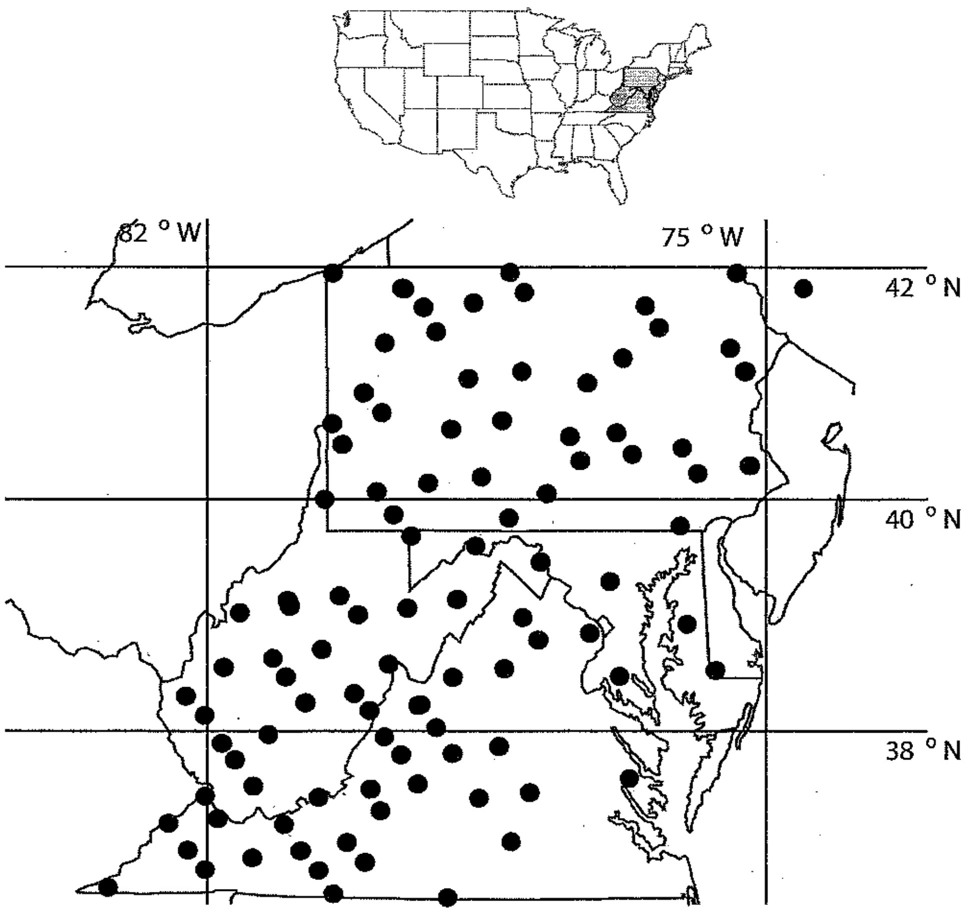


Figure 1. Mid-Atlantic Region of the U.S. With the Wadable Stream Sampling Sites Used in This Study.

102 stream segments in first to third (Strahler) order wadable streams (Figure 1). These segments were selected for sampling using a spatially balanced probability design (Stevens, 1997; Stevens and Olsen, 1999). Inclusion probabilities for each sampled stream segment were determined using the sample sizes for each Strahler order and the total length of streams within each order in the region. Sampling locations within stream segments were chosen randomly. Quantitative data for stream macroinvertebrates, habitat, and water quality were collected at each site (for specifics see Lazorchak *et al.*, 1998; Kaufman and Robinson, 1998; Herlihy *et al.*, 2000; and Klemm *et al.*, 2002). Sampling took place during a two-month sampling window each year from April through mid-June.

Stream Macroinvertebrate Data

Stream benthic macroinvertebrates are a robust measure of stream condition, integrating temporal pollutant exposure. They are responsive to changes in in-stream sediment levels (Davis and Lathrop, 1992; Covich, 1999). Benthic stream community taxa in the orders Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies) (collectively known as EPT) are considered reasonably sensitive indicator organisms since they exhibit a decrease in taxa richness with increased degradation of stream conditions (Loch *et al.*, 1996; Barbour *et al.*, 1999; Zweig and Rabeni, 2001; Klemm *et al.*, 2002; Kaller and Hartman, 2004). The EPT taxa were used to identify impacted stream segments in the mid-Atlantic; when EPT taxa were less than 9 in the stream segments, the stream segments were considered to be impacted (Davis and Scott, 2000; Klemm *et al.*, 2002).

Indicator for Sedimentation

Sediments, including suspended and bedded, can directly affect stream biota or indirectly affect stream biota through changes in habitat. For example, excessive suspended sediments in aquatic systems can cause increased turbidity and decreased light penetration. Altered light regimes can directly alter primary productivity and increase shading of submerged macrophytes (Canfield *et al.*, 1985; Best *et al.*, 2001). Excess fine sediments can fill in gaps between larger substrate particles, embedding the larger particles and eliminating interstitial spaces that would otherwise be used as habitat for reproduction, feeding, and cover for invertebrates and fish (Suttle *et al.*, 2004). For example, bedded sediments in streams and rivers can cause the loss of spawning habitat for salmonids due to increased embeddedness (Young *et al.*, 1991).

For the purpose of this paper, percent fines in the substrate is used as a surrogate indicator for sedimentation in streams. The percent fines (silt/clay fraction, less than 0.06 mm) is a direct measure of the smallest class of sediments and is strongly correlated with sediment embeddedness, a source of the most likely to be resuspended sediment, and an indirect measure of suspended sediment levels in the water column. Streams containing a larger fraction of fine sediment would be expected to have a benthic community at greater risk for impact (Zweig and Rabeni, 2001). Details of the protocols for sample collection and analysis for percent fines can be found in Lazorchak *et al.* (1998), Kaufmann *et al.* (1999), and Klemm *et al.* (2002). Percent fines is determined by visual examination at 11 equally spaced stream transects. The data used were restricted to stream segments with pools.

Reference Conditions

Reference conditions are expectations as to the condition of biological communities in the absence of any human disturbance (Plafkin *et al.*, 1989; Gerritsen *et al.*, 1994). These conditions provide an estimate of natural variability in biological condition and habitat quality that can be expected to occur. Few streams in the mid-Atlantic area are undisturbed. Reference conditions in the mid-Atlantic have been identified (Waite *et al.*, 2000) using a set of selected chemical and habitat conditions from unimpacted or minimally impacted streams in the area. These chemical and habitat measures are used to identify the best available biological conditions in the area streams.

The chemical and habitat parameters used in the selection of reference conditions for wadable streams in the mid-Atlantic area (Waite *et al.*, 2000) are: acid neutralizing capacity (> 50 μeq/L), chloride (< 100 μeq/l), sulfate (< 400 μeq/l), total nitrogen (< 750 μg/l), total phosphorous (< 20 μg/l), and mean rapid bioassessment protocol (RBP) habitat score (> 15). The RBP habitat score encompasses the variety and quality of the substrate, channel morphology, bank structure, and riparian vegetation. It is measured on a scale of 1 to 20, where 1 is very poor habitat and 20 is excellent habitat (Barbour *et al.*, 1999). Waite *et al.* (2000) define a stream segment as reference if all six of the chemical and habitat parameters are within the desired levels. Reference sites were identified that met these criteria for the extant data. Thus, the number and distribution of reference site conditions were not predetermined and were used strictly for comparative purposes.

METHODS

Conditional Probability Analysis (CPA)

The probability of observing a certain event y is denoted P(y). Data acquired with a probability survey design provide estimates of the probability of occurrence for a sampled variable. For example, consider a sampling frame that includes all stream segments in a state. If 75 percent of the sampled stream segments exhibit impacted benthic communities, then the likelihood of observing benthic impact in any of the stream segments in the state is 75 percent.

For use in developing a numeric water quality criterion, a conditional probability statement provides the likelihood (probability) of impact, if the value of a pollution metric (threshold) is exceeded. Conditional probability is the probability of an event when it is known that some other event has occurred, and is denoted

$$P(y \mid x^*) \tag{1}$$

where y is the event of interest and x^* is the other event that has occurred previously. For criterion development, x^* is replaced with $x > x_C$, where x_C is the specific threshold that is exceeded. Therefore, the conditional probability statement in this paper is the probability of an event y occurring, when it is known that some event x has occurred and x has exceeded some threshold x_C [$P(y \mid x > x_C)$]. For water quality criteria, this implicitly assumes that as the pollution metric of interest increases, the likelihood of an impact on biological condition increases. For responses that increase as the pollution metric decreases (e.g., dissolved oxygen concentration), the conditional probability description is reversed [$P(y \mid x < x_C)$].

The pollution metric, x, is treated as a discrete random variable. The probability mass function (pmf) of x (the probability that x is a specific value) is represented by $f_i(x_i)$, where i is 1,2, ... N, and N is the total number of samples. The inclusion probability provides the probability for selection of a given sample, that it can be different for each sample, and is used as a weighting factor for statistical estimation. Every sampling unit has a nonzero probability of being selected. The sum of f_i over all possible sample sites is 1.

It is assumed that an appropriate response variable for stream condition, y_i, exists, which also is a discrete random variable, and equals 0 for good conditions and 1 for impacted conditions. The y_i will be paired with values of the pollution metric, x_i. The value of y_i will be dichotomous (0,1) depending on the value of x_i and can be written as a function of x_i, $y_i = g(x_i)$. The conditional probability, in this work, of impacted conditions if specific value, x_C, is exceeded, is

$$P(\mathbf{y}=1 \mid \mathbf{x} > \mathbf{x}_C) = \frac{P(\mathbf{y}=1, \mathbf{x} > \mathbf{x}_C)}{P(\mathbf{x} > \mathbf{x}_C)} \tag{2}$$

where $P(y = 1, x > x_C)$ is probability of joint occurrence for the two events. This equation is the definition of conditional probability (Hogg and Ledolter, 1992). The conditional probability can be expressed as

$$P(\mathbf{y}=1 \mid \mathbf{x} > \mathbf{x}_C) = \frac{\sum_{x_i > x_c} g(x_i) f_i(x_i)}{\sum_{x_i > x_c} f_i(x_i)} \tag{3}$$

Functionally, to determine the probability of impact in the stream when some value of the pollution metric, x_C, is exceeded, $P(y = 1 \mid x > x_C)$, a two-step procedure is used. Using the survey data, one identifies a subset of the sampled resource (e.g., stream segments) for which $x > x_C$ (i.e., the stream segments are stratified based on the value of the pollution metric x); and from this subset of stream segments in which $x > x_C$, one determines those segments in which the biological conditions are impacted. This is the subset of the sampled stream segments in which the pollution metric exceeds a specific value (x_C), and which are also biologically impacted.

The probability of the biology being impacted in stream segments when $x > x_C$ over the entire range of observed x provides an empirical conditional probability curve. Confidence intervals (CIs) for this empirical curve can be estimated by bootstrap resampling (Manly, 1997). Bootstrapping assumes that the distribution of a population can be determined by resampling the original data. A bootstrap sample consists of drawing a sample of size N from the original data (of size N) with replacement, which is then used to calculate a bootstrap value for conditional probability, $P(y = 1 \mid x > x_C)$. One thousand samples were generated for the bootstrap distribution. The 90 percent and 95 percent CIs were determined from the empirical percentiles (Insightful Corp., 2001).

Identifying Thresholds of Impact

Threshold levels for pollutants that elicit different levels of biological impact in stream segments of a region need to be identified for eventual use in developing criteria. A threshold of impact was identified as a changepoint separating the empirical conditional probability curve into two parts, that part of the curve above the changepoint and that which is below it. For those samples that are above the changepoint, the probability of impact is different from what one would expect for the entire geographic area. A confounding factor in the identification of a changepoint is that these two groups created by the changepoint are not independent (i.e., the numbers used to create the points above the changepoint are a subset of the numbers used to create the points below the changepoint). Thus, a traditional t-test cannot be used in the determination of the changepoint since the data are not independent (Venables and Ripley, 1997). The identification of the changepoint was by using a weight-of-evidence approach with three different techniques. These techniques are: nonoverlapping confidence intervals, change in curvature of fitted curve, and nonparametric deviance reduction. Other possible techniques could be used to identify a changepoint. In this demonstration, specific values for factors and CIs were selected only as examples. Values used in an actual application of this approach would depend on the particular management requirements and objectives.

The use of nonoverlapping CIs to determine a changepoint involves determining when the lower CI of the empirical curve no longer overlaps the upper CI of the unconditional value (Cherry, 1996; Rahlfs, 1997; Cherry, 1998; Austin and Hux, 2002). This procedure is a conservative estimate for significant difference, since the CIs could overlap when the values are significantly different (Austin and Hux, 2002). The bootstrap percentile CIs based on a bootstrap distribution of 1,000 samples were used for this evaluation. The α-level for the nonoverlapping CI must be adjusted to account for the one-sided nature of this test, whereas the α-level for developing the CIs for the curves was based on a two-sided test (i.e., a factor of 2 in the α-level).

The second technique used for selecting a threshold of impact through changepoint identification was to fit an equation to the empirical curve for conditional probability. The following constraints were used: the conditional probability approaches the unconditional value, $P(y = 1)$, as x goes to the minimum x-value; the conditional probability approaches 1 as x goes to the maximum value; and there is a curvature change at the inflection point of the curve. The following functional form satisfies these constraints

$$P(\mathbf{y}=1 \mid x > x_C) = \begin{cases} \mathbf{1+(D_0-1)/(1+exp(B_0(x_C-x_0))), for\ x_C > x_0} \\ \mathbf{1+(D_0-1)/(1+exp(B_1(x_C-x_0))), for\ x_C \le x_0} \end{cases} \quad (4)$$

where exp is the exponential function to base e, D_0 is unconditional probability value $P(y = 1)$, x_0 is the changepoint where curvature changes, B_0 is curvature for values of $x_C > x_0$, and B_1 is curvature for values of $x_C \le x_0$. The parameters x_0, B_0, and B_1 are determined from a nonlinear least squares regression (Venables and Ripley, 1997). Uncertainty in the parameters are estimated from the standard errors generated by the regression software and, where possible, by computing asymmetric confidence intervals (Venables and Ripley, 1997). The residuals from the regression were checked for normality. While it may be generally possible to fit Equation (4) to the empirical curve, the curvature values (B_0 and B_1) may not be significantly different, and a threshold would not be identified with this technique.

The third technique uses nonparametric deviance reduction to determine the changepoint. This approach determines the dividing point for splitting the data into two groups, resulting in the largest reduction in the deviance in the data (Qian *et al.*, 2003). The deviance is defined as

$$D = \sum_{i=1}^{N} (P_i - P^*)^2 \quad (5)$$

where D is the deviance, N is the sample size, Pi is the conditional probability $P(y = 1 \mid x > x_i)$, and P^* is the mean of P_i based on a sample size of N. When the data are divided into two groups, the sum of the deviance for the two subgroups is always less than or equal to the deviance for the entire data set. When the split in the data minimizes the deviance, the threshold is identified. This approach has been used to detect ecological changes along an environmental gradient (Qian *et al.*, 2003). Qian *et al.* (2003) compared results of deviance reduction with a Bayesian hierarchical modeling approach and found that the nonparametric approach provides similar results with the Bayesian analysis.

The deviance reduction point generally can be determined, but it may or may not be of biological significance. Uncertainty in the deviance reduction changepoint (90 percent and 95 percent CIs) is estimated from the empirical percentiles for the bootstrap distribution from resampling 1,000 times (Manly, 1997). An approximate χ^2 test was used to determine the significance of the changepoint. The test assumes that the deviance reduction divided by the scale parameter is approximately χ^2 distributed with 1 degree of freedom (Venables and Ripley 1997). A large deviance reduction will result in a small p-value, and the consequent rejection of the null hypothesis (H_0: no changepoint).

Biological Importance of Identified Thresholds

For use in criteria development, some level of biological importance needs to be associated with the threshold of impact value that is identified. The changepoint value determined by each technique must separate the samples so that the probability of impact for samples above the threshold would be different from what one would expect for the entire geographic area. A summary of literature values on the response of fish and benthic invertebrates at low reported levels of percent fines in the substrate (Newcombe and Jensen, 1996; Bash *et al.,* 2001; Berry *et al.,* 2003) was used to identify biological importance.

Statistical Analysis of Data

The cumulative distribution function (CDF), the conditional cumulative distribution function (CCDF), and their reverses were used to supplement the conditional probability analysis. The CDF gives probability that x is less than or equal to x_C

$$P(x \le x_C) \approx F(x_C) = \sum_{x_i \le x_c} f_i(x_i) \tag{6}$$

The reverse CDF is the probability that x is greater than x_C, which is the complement of Equation (6), or

$$P(x > x_C) \approx 1 - F(x_C) = 1 - \sum_{x_i \le x_c} f_i(x_i) = \sum_{x_i > x_c} f_i(x_i) \tag{7}$$

The CCDF is the distribution for a subset of the total data, subsetted by (or conditioned on) a second variable [F(y | x)]. The reverse CCDF is similar to Equation (7), that is, 1-F(y | x). The reverse functions are consistent with the CPA results, which are expressed as a threshold (i.e., exceeding some value x_C).

RESULTS

The CDF for EPT taxa richness is shown in Figure 2. Approximately 42 percent of the stream miles across the region were observed to have EPT taxa

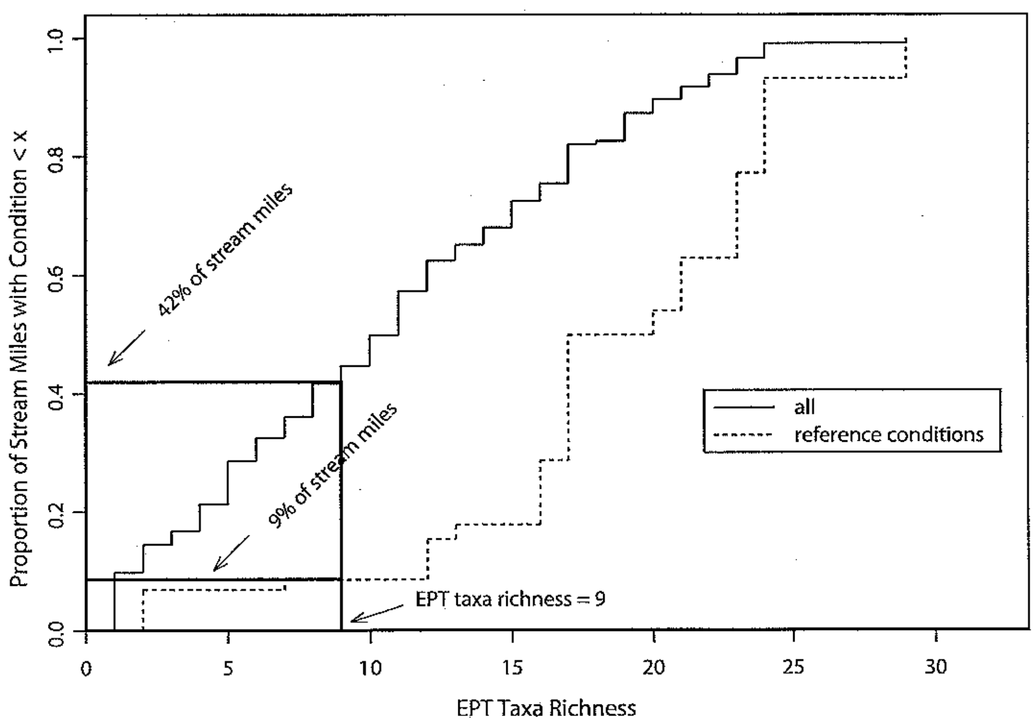


Figure 2. Cumulative Distribution Function for EPT Taxa Richness for All Stream Miles and for Stream Miles That Exhibit Reference Condition Characteristics.

richness less than 9, indicating impacted benthic community conditions. Out of 100 stream segments that had valid values for the indicators used in this study, 16 met the reference condition requirements. In 91 percent of the reference condition stream miles, benthic communities were found that had EPT taxa equal to or greater than 9 (Figure 2). The EPT taxa richness generally declines as the percent fines increases (Figure 3, correlation coefficient, r, is -0.50). The fraction of EPT taxa richness variance explained using a linear regression with percent fines as the predictor is 0.25, suggesting that percent fines does appear to have a substantial effect on EPT taxa richness.

The reverse CDF and reverse CCDFs for percent fines in the substrate were expressed as a proportion of stream miles (Figure 4). The sampled stream segment values were weighted by inclusion probabilities to convert to stream miles. The distribution for impacted benthic communities is displaced to the right of that for benthic communities in good condition (Figure 4). The distribution for reference conditions is shifted to the left (towards lower percent fines) than that for unimpacted streams (Figure 4), since these are the best observed conditions.

The CPA approach suggests that when percent fines in the substrate is greater than 49 percent, there is a 100 percent probability that the benthic communities are impacted (Figure 5). All sites with percent fines in the substrate in excess of 49 percent had EPT taxa richness less than 9. As the percent fines approaches zero, there is a background level of impact on EPT taxa richness from all sources of stress in the region (mean = 42 percent, 95 percent confidence interval of 30 to 56 percent). Thus, irrespective of the level of percent fines in the substrate, approximately 42 percent of the stream miles in the region will likely exhibit an impact on EPT taxa richness. Therefore, to detect a significant signal due to percent fines in the substrate affecting the EPT taxa richness, the upper confidence limit on the estimate of the background impact (e.g., 56 percent, Figure 5) must not overlap with the lower confidence limit on the probability of benthic impact curve in Figure 5. The point at which this occurs is when the percent fines in the substrate is 14.8 percent (Figure 5). This is a threshold of impact, and is statistically distinguishable from background within this geographic area. The mean probability of observing impacted EPT taxa richness associated with this threshold is 67 percent.

The CPA identified threshold of 14.8 percent fines (from nonoverlapping CIs) would translate into approximately 47 percent of the total stream miles in the geographic area exceeding the threshold (from Figure 4). Similarly, only a small percentage of streams with reference condition characteristics (6 percent) or good benthic conditions (21 percent)

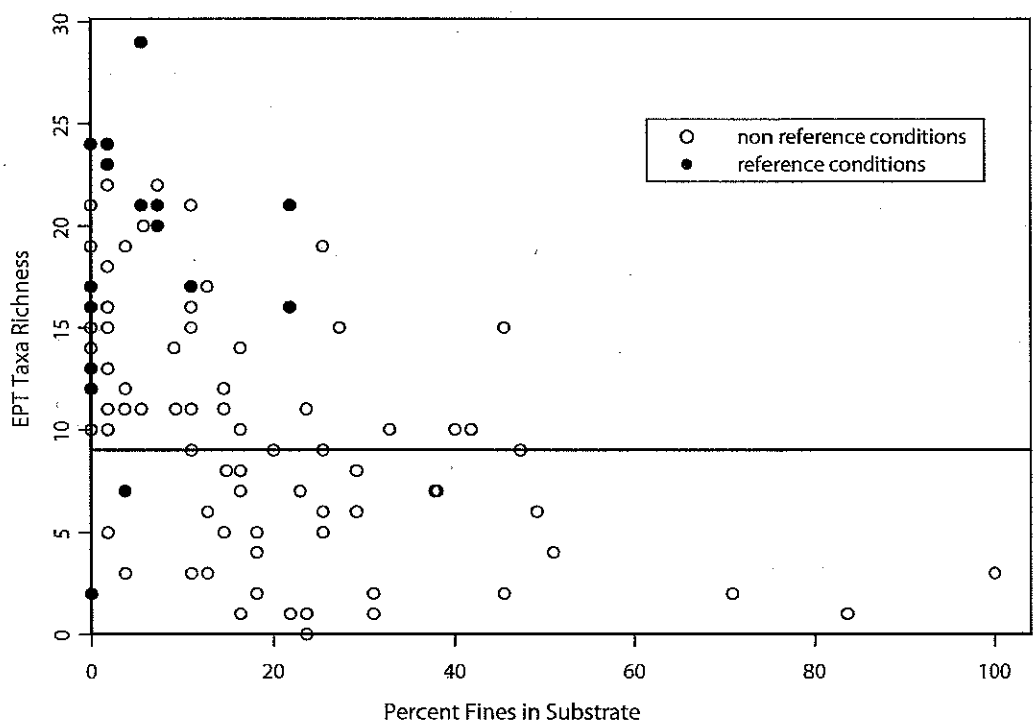


Figure 3. Plot of EPT Taxa Richness Against Percent Fines in Substrate (silt/clay fraction, less than 0.06 mm). Horizontal line for EPT taxa richness = 9. Solid circles are stream segments that exhibit reference condition characteristics. Open circles are segments not satisfying reference condition characteristics.

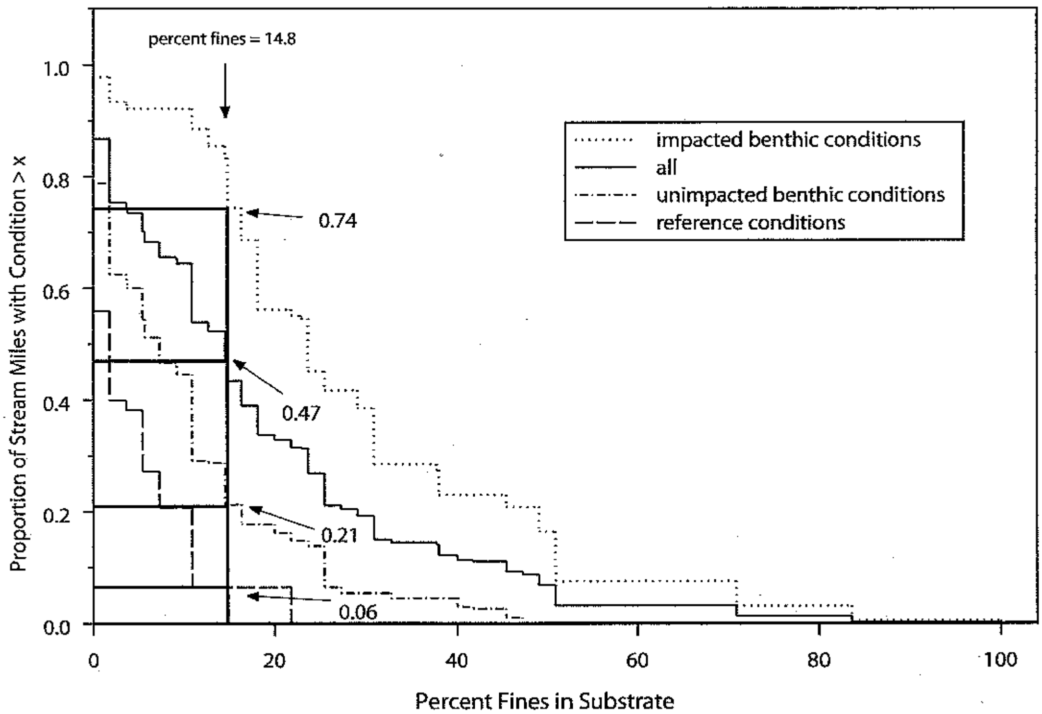


Figure 4. Reverse Cumulative Distribution Function (CDF) for Percent Fines in the Substrate (silt/clay fraction, less than 0.06 mm) for Stream Miles Across Entire Area, and Reverse Conditional CDFs of Stream Miles for Impacted Benthic Conditions (EPT taxa richness less than 9), Unimpacted Benthic Conditions (EPT taxa richness equal to or greater than 9), and Reference Conditions. Horizontal lines are where the threshold of 14.8 percent fines intersects curves.

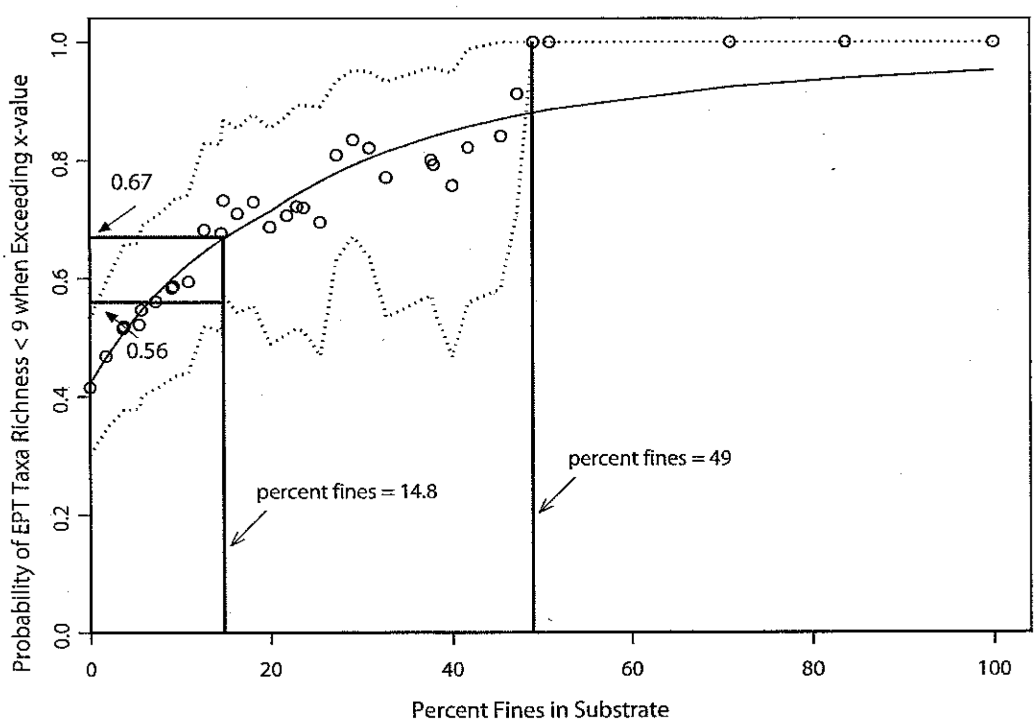


Figure 5. Probability of Observing EPT Taxa Richness Less Than 9 (benthic impact) in Mid-Atlantic Streams (open circles) if Specified Value of Percent Fines in the Substrate (silt/clay fraction, less than 0.06 mm) is Exceeded. Solid line is fit of Equation (4) (see Table 1). Dotted lines are 95 percent confidence intervals (CIs) from bootstrap estimation.

would exceed the 14.8 percent fines threshold, but a much larger percentage of impacted streams (74 percent) would exceed it (from Figure 4). These values provide an estimate of the number of "false positives" for this value of a threshold for percent fines as the indicator of sedimentation. Because multiple stressors often impact stream communities, one cannot estimate the "false negatives." A community not stressed by the stressor of interest might be stressed in some other way.

The coefficients from the nonlinear least squares regression for Equation (4) are given in Table 1, with the fitted curve shown in Figure 5. This technique also determined a threshold of impact of 14.8 percent fines in the substrate (Table 2). Using the nonparametric deviance reduction technique, a threshold of impact of 15.3 percent fines in the sediment, with p = 0.03, was identified. All three techniques for identifying a threshold of impact from the conditional probability analysis yielded consistent results (Table 2).

TABLE 1. Coefficients (mean value and confidence limits) From Nonlinear Least Squares Regression of Equation (4) for Percent Fines in Substrate Against Probability of Impacted Benthic Community (EPT taxa richness less than 9).

	Mean	90 Percent CI	95 Percent CI
B_0	-0.0328	(-0.049, -0.03)	(-0.0516, -0.0292)
B_1	-0.159	(-0.194, -0.138)	(-0.202, -0.133)
x_0	14.8	(13.4,15.9)	(13.1, 16.2)

These three different techniques all determined thresholds of impact that separated the data such that a difference could be detected from what would be expected for the entire geographic area. For the first technique, the existence of nonoverlapping CIs provided the difference. For the second technique, the nonoverlap of CIs for the curvature parameters established the difference. In the third technique, the null hypothesis of no changepoint was rejected (p = 0.03).

The literature supports a biological response of fishes to the thresholds for percent fines in the substrate: survival of salmonids have been shown to be negatively affected when fines exceed 10 to 20 percent (McNeil and Ahnell, 1964; Burns, 1970; Tappel and Bjornn, 1983; Chapman, 1988; Peterson *et al.* 1992; Argent and Flebbe, 1999). These studies for salmonids were all for a larger sediment size range (silt/clay/sand) than was chosen for purposes of this demonstration. However, the silt/clay fraction is always less than or equal to the silt/clay/sand fraction of the same sample. These threshold values are consistent with the lower end of reported response levels in the literature (see Newcombe and Jensen, 1996).

DISCUSSION

Scientifically defensible numeric criteria are highly desirable for water quality protection programs responsible for preventing the impairment of aquatic systems. Historically, for single chemical pollutants, carefully controlled laboratory bioassays have been conducted. From the dose-response relationship derived from these bioassays, an appropriate criterion for the pollutant was developed that is protective of aquatic life. Unfortunately, this historical approach is not applicable for nonpoint source pollution. At low levels, nonpoint source materials may not be a pollutant, but may be necessary for the functioning of the aquatic systems (e.g., nutrients, sediments) and their levels may naturally fluctuate over a geographic area. Only when the levels become excessive (usually in conjunction with anthropogenic activity), do they

TABLE 2. Summary of Thresholds of Impact for Percent Fines in Substract (silt/clay fraction, less than 0.06 mm) Identified by Conditional Probability Analysis (mean threshold and confidence intervals, when available) Using Three Techniques.

	Threshold (percent)	90 Percent CI (percent)	95 Percent CI (percent)
Nonoverlapping Confidence Intervals	12.7[1] 14.8[2]		
Change in Curvature of Fitted Curve	14.8	13.4 to 15.9	13.1 to 16.2
Nonparametric Deviance Reduction	15.3 (p = 0.03)	11.8 to 26.4	8.9 to 26.4

[1]From two-sided 90 percent CI.
[2]From two-sided 95 percent CI.

become pollutants and require criteria development for their control. The development of scientifically defensible approaches for establishing thresholds, and eventually criteria, for nonpoint source pollution is a critical need for regulatory agencies (USEPA, 2003b). Any approach undertaken to develop nonpoint source criteria must take into account the natural variability of the pollution occurring across the geographic area of interest, and the impact of other stressors that are likely impacting the aquatic systems as well.

Use of the CPA can establish realistic thresholds for the impact on stream biotic condition by nonpoint source pollution. This approach was applied to establish a threshold of sediment impact on a susceptible biological community in wadable streams in the mid-Atlantic region of the U.S. The mid-Atlantic was selected because of the extensive amount of research and monitoring of streams in this region (e.g., Boward *et al.,* 1999; USEPA, 2000b), which provided the information base needed that would satisfy the conditions for application of CPA. The necessary conditions were: (1) monitored data must be collected based on a probability based sampling design; (2) there must be some metric that can quantify the pollution parameter of interest; (3) there must be a response metric sufficiently sensitive to respond to the extant levels of the pollution parameter of interest; (4) independent studies must be available that identify the characteristics of an impacted response metric; and (5) the pollution parameter must be capable of exerting a strong effect on the response metric.

The streams in the mid-Atlantic region met these criteria. Sufficient empirical data were available from a probability monitoring design for wadable streams in the region (see McDonald *et al.,* 2004; data are available from http://www.epa.gov/emap/). The probabilistic sampling allows for statistically rigorous extrapolation from the sites sampled to the entire region of interest. Sedimentation was a major stressor in these streams (Boward *et al.,* 1999; USEPA, 2000b), and sufficient information was available on the percentage of fines in the substrate to allow its use as a surrogate for sedimentation (Klemm *et al.,* 2002). Data had been collected on EPT taxa richness at these sites and related to stream condition (USEPA, 2000b). Davis and Scott (2000) had determined a level of EPT taxa species richness (less than 9) below which mid-Atlantic highland streams were likely to be impacted and in relatively poor condition. Last, EPT taxa richness responded strongly to sedimentation (Figure 3).

It was decided not to develop thresholds for protecting against impact based on not exceeding a pollution metric value (i.e., $P(y = 0 \mid x \leq x_C)$, where y = 0 represents unimpacted conditions). The approach of looking for a threshold which, if not exceeded, would indicate a high probability of encountering unimpacted conditions, is the approach taken for criteria developed from laboratory toxicological studies. What makes this appropriate for laboratory studies is the ability to control for all stressors other than the one for which criteria are being developed. These studies provide y = 0 (unimpacted) if $x < x_C$. This approach does not work when dealing with actual field data, as one cannot control for all of the myriad stressors that are affecting the biological communities. The biological response observed reflects the cumulative response to all of the stressors. As the data for a specific stressor is analyzed, and as the magnitude of that stressor decreases, one would not expect a continual increase in the likelihood of unimpacted conditions, unless all of the stressors are strongly correlated. Reducing one stressor would still leave other stressors eliciting impact (in the case of 42 percent of the stream miles in Figure 2). Therefore, any calculation of $P(y = 0 \mid x \leq x_C)$ would likely be confounded by other stressors (i.e., value for unconditional probability would not be zero). Thus, the thresholds identified with the conditional probability analysis (and based on the reverse CCDF) are thresholds of impact, above which the likelihood of impact is high. One is able to pull out the signal in the mixture of multiple stressors with the conditional probability analysis because the stressor chosen was strong enough to elicit an impact as the magnitude of the stressor increased. With the conditional probability analysis approach, one is protecting the aquatic resource against the likelihood of impact.

The EPT taxa richness metric was used to identify impacted stream communities. The taxa in the orders Ephemeroptera (mayflies), Plecoptera (stoneflies), and Trichoptera (caddisflies) respond similarly to sedimentation as estimated by percent fines in the substrate (Figure 3), with the probability of impact increasing as percent fines increased. Other functional biological groupings that responded similarly to EPT were benthic invertebrate scrapers and intolerant taxa richness, while noninsect benthic invertebrates, benthic invertebrate scavengers, and tolerant taxa richness responded with the probability of impact decreasing as percent fines increased. Conditional probability plots could have been generated using any of these groupings of the benthic invertebrate community, if a level of biological impact could be independently assigned (similar to EPT taxa richness less than 9).

With the CPA approach, traditional statistics cannot be used to ascertain the threshold of impact. Instead, a weight-of-evidence approach based on three separate techniques was used. However, this assumed that there was a consistency in the threshold levels identified using these disparate techniques. The CPA does provide relatively consistent thresholds of impact for the percent fines in the substrate, irrespective of which of the three techniques were applied (Table 2). These CPA thresholds contrasted markedly with the threshold values obtained with the two *ad hoc* approaches currently practiced for developing aquatic criteria based on monitoring data from sites across a geographic area (USEPA, 2000a). The two techniques consist of setting thresholds with either the levels of stressor associated with streams in the 75th percentile of the reference stream miles sampled or the 25th percentile of all stream miles sampled. Using these approaches, the threshold for percent fines would be approximately 1.9 percent based on all stream miles and approximately 7.1 percent based on reference stream miles (Figure 4). These values are substantially lower than estimates from this study, and fall outside of the 95 percent confidence limits (Table 2). These values are also substantially lower than literature thresholds for percent fines.

While the literature supports the biological importance associated with the thresholds identified, the agreement of these threshold values with lower values from the literature does not validate the thresholds. However, it does gives credence to the conditional probability analysis approach for identifying realistic thresholds for use in development of criteria for protection of aquatic life. The CPA can provide environmental managers with an additional tool to evaluate the tradeoffs of setting different criteria. Using CPA, environmental managers can examine a given criterion and the tradeoffs: the likely number of stream miles that actually have good biological communities when the criterion level is exceeded and the number of streams that have impacted biology when the criterion level is not exceeded. This would allow the protection of the ecosystems to be more quantitative and explicit when being weighed in conjunction with economic considerations.

The CPA approach for threshold of impact can be combined with information on reference conditions and toxicological data to develop candidate values for water quality criteria. The steps in this process that provide the candidate values are listed below.

1. Acquire the survey data (probability based) that includes candidate pollutant for criterion development and biological response metrics.
2. Use available information on reference conditions (physical, chemical, and habitat metrics) to define impacted biological conditions in terms of biological metrics.
3. Conduct conditional probability analysis (probability of impact if value of candidate pollutant is exceeded).
4. Identify threshold of impact from conditional probability analysis results.
5. Evaluate identified threshold of impact against reference conditions, impacted conditions, and good conditions. Evaluate identified threshold for biological importance.

Nonetheless, additional work must be done to evaluate and validate the conditional probability analysis approach for identification of thresholds of impact. This could be accomplished by using other survey data from other geographic areas where the conditions for CPA are met. The approach should also be tested with other pollution and response parameters to confirm that this is a robust approach, which can be used for identifying realistic thresholds of impact.

SUMMARY

The conditional probability analysis approach can be used to develop realistic thresholds of impact for nonpoint source pollution on aquatic benthic communities in waterbodies across a region. However, this approach is predicated on the following conditions: (1) monitored data have been collected based on a probability based sampling design; (2) some metric must be available that can quantify the pollution parameter of interest; (3) a response metric sufficiently sensitive to respond to the extant levels of the pollution parameter of interest must be available; (4) independent studies must be available that identify the characteristics of an impacted response metric; and (5) the pollution parameter must be capable of exerting a strong effect on the response metric. In the example presented here, realistic thresholds of impact on EPT taxa richness due to sedimentation in mid-Atlantic wadable streams were identified. Threshold values from CPA were found to be in the range of 12 to 15 percent fines in the substrate, based on three different techniques for threshold identification. These threshold values were found to be consistent with existing literature from laboratory and field studies. These values were quite different from those determined with the current *ad hoc* practice, with the current practice appearing to produce overly prescriptive thresholds. Development of scientifically defensible thresholds are a necessary first step for managers in establishing protective criterion. However, thresholds determined with CPA, or other methods, should not be used exclusively to set water quality criteria, as other additional factors (e.g., designated uses, ecotoxicological data, economics) must be considered by managers when establishing criteria and standards.

ACKNOWLEDGMENTS

Thanks to Brian Hill, Jim Wickham, Walter Berry, Jerry Pesch, John Van Sickle, Phil Kaufman, and anonymous reviewers for the critical and constructive reviews that they provided on various versions of this manuscript. Special thanks goes to Steve Hedtke for the encouragement to apply the conditional probability analysis approach to sediment criteria and to the EMAP-Surface Waters team for generating high quality data and making these data available to all. S-Plus software was used for statistical analyses and graphical displays. The research described in this paper was funded by the U.S. Environmental Protection Agency. This paper was not subjected to Agency review, and therefore does not necessarily reflect the views of the Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

LITERATURE CITED

Argent, D.G. and P.A. Flebbe, 1999. Fine Sediment Effects on Brook Trout Eggs in Laboratory Streams. Fisheries Research 39(3): 253-262.

Austin, P.C. and J.E. Hux, 2002. A Brief Note on Overlapping Confidence Intervals. Journal of Vascular Surgery 36(1):194-195.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling, 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish (Second Edition). EPA 841-B-99-002, U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Bash, J., C. Berman, and S. Bolton, 2001. Effects of Turbidity and Suspended Solids on Salmonids. University of Washington, Center for Streamside Studies. *Available at* http://depts.washington.edu/cwws/Outreach/Publications/Salmon%20and%20Turbidity.pdf. *Accessed in* August 2005.

Berry, W., N. Rubinstein, B. Melzian, and B. Hill, 2003. The Biological Effects of Suspended and Bedded Sediment (SABS) in Aquatic Systems: A Review. U.S. Environmental Protection Agency, Office of Research and Development, Internal Report. *Available at* http://www.epa.gov/waterscience/criteria/sediment/appendix1.pdf. *Accessed in* August 2005.

Best, E.P.H., C.P. Buzzelli, S.M. Bartell, R.L. Wetzel, W.A. Boyd, R.D. Doyle, and K.R. Campbell, 2001. Modeling Submersed Macrophyte Growth in Relation to Underwater Light Climate: Modeling Approaches and Application Potential. Hydrobiologia 444(1-3): 43-70.

Borja, A., V. Valencia, J. Franco, I. Muxika, J. Bald, M. J. Belzunce, and O. Solaun, 2004. The Water Framework Directive: Water Alone, or in Association With Sediment and Biota, in Determining Quality Standards? Marine Pollution Bulletin 49(1-2):8-11.

Boward, D.M., P.F. Kazyak, S.A. Stranko, M.K. Hurd, and T.P. Prochaska, 1999. From the Mountains to the Sea: The State of Maryland's Freshwater Streams. Maryland Department of Natural Resources, Monitoring and Non-tidal Assessment Division, EPA 903-R-99-023, Annapolis, Maryland.

Burns, J.W., 1970. Spawning Bed Sedimentation Studies in Northern California Streams. California Fish and Game Quarterly 56(4):253-270.

Canfield, D.E.G., K.A. Langeland, S.B. Linda, and T.W. Haller, 1985. Relations Between Water Transparency and Maximum Depth of Macrophyte Colonization in Lakes. Journal of Aquatic Plant Management 23:25-28.

Chapman, D.W., 1988. Critical Review of Variables Used to Define Effects of Fines in Redds of Large Salmonids. Transactions of the American Fisheries Society 117(1):1-21.

Cherry, S., 1996. A Comparison of Confidence Interval Methods for Habitat Use-Availability Studies. Journal of Wildlife Management 60(3):653-658.

Cherry, S., 1998. Statistical Tests in Publications of the Wildlife Society. Wildlife Society Bulletin 26(4):947-953.

Covich, A.P., 1999. The Role of Benthic Invertebrate Species in Freshwater Ecosystems. Bioscience 49(2):119-127.

Davis, W.R. and J. Scott, 2000. Mid-Atlantic Highlands Streams Assessment: Technical Support Document. EPA/903/B-00/004, Mid-Atlantic Integrated Assessment Program, Region 3, U.S. Environmental Protection Agency, Ft. Meade, Maryland.

Davis, W.S. and J.E. Lathrop, 1992. Freshwater Benthic Macroinvertebrate Community Structure and Function. *In:* Sediment Classification Methods Compendium, EPA 823-R-92-006, pp. 8-1 to 8-26. U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

Fisher, D.J. and D.T. Burton, 2003. Comparison of Two U.S. Environmental Protection Agency Species Sensitivity Distribution Methods for Calculating Ecological Risk Criteria. Human and Ecological Risk Assessment 9(3):675-690.

Gerritsen, J., J. Green, and R. Preston, 1994. Establishment of Regional Reference Conditions for Stream Biological Assessment and Watershed Management. *In:* Proceedings, Watershed '93, A National Conference on Watershed Management. EPA/840/R-90/002, U.S. Environmental Protection Agency, Office of Water, Washington, D.C., pp. 797-801.

Herlihy, A.T., D.P. Larsen, S.G. Paulsen, N.S. Urquhat, and B.J. Rosenbaum, 2000. Designing a Spatially Balanced, Randomized Site Selection Process for Regional Stream Surveys: The EMAP Mid-Atlantic Pilot Study. Environmental Monitoring and Assessment 63(1):95-113.

Hogg, R.V. and J. Ledolter, 1992. Applied Statistics for Engineers and Physical Scientists. Macmillan Publishing Co., New York, New York.

Hohreiter, D.W., and D.K. Rigg, 2001. Derivation of Ambient Water Quality Criteria for Formaldehyde. Chemosphere 45(4-5):471-486.

Insightful Corp., 2001. S-Plus 6 for Windows, Guide to Statistics, Volume 2. Insightful Corp., Seattle, Washington.

Jimenez, B., J. Ramos, and L. Quezada, 1999. Analysis of Water Quality Criteria in Mexico. Water Science and Technology 40(10):169-175.

Kaller, M.D., and K.J. Hartman, 2004. Evidence of a Threshold Level of Fine Sediment Accumulation for Altering Benthic Macroinvertebrate Communities. Hydrobiologia 518:95-104.

Kamizoulis, G. and L. Saliba., 2004. Development of Coastal Recreational Water Quality Standards in the Mediterranean. Environment International 30(6):841-854.

Kaufmann, P.R., P. Levine, E.G. Robinson, C. Seeliger, and D.V. Peck, 1999. Quantifying Physical Habitat in Wadeable Streams. EPA/620/R-99/003, U.S. Environmental Protection Agency, Washington, D.C.

Kaufmann, P.R. and E.G. Robinson, 1998. Physical Habitat Characterization in Environmental Monitoring and Assessment Program – Surface Waters. *In:* Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams, J.M. Lazorchak, D.J. Klemm and D.V. Peck (Editors). EPA/620/R-94/004F, U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C., pp. 77-118.

Kay, D., J. Bartram, A. Pruss, N. Ashbolt, M.D. Wyer, J.M. Fleisher, L. Fewtrell, A. Rogers, and G. Rees, 2004. Derivation of Numerical Values for the World Health Organization Guidelines for Recreational Waters. Water Research 38(5):1296-1304.

Klemm, D.J., K.A. Blocksom, W.T. Thoeny, F.A. Fulk, A.T. Herlihy, P.R. Kaufmann, and S.M. Cormier, 2002. Methods Development and Use of Macroinvertebrates as Indicators of Ecological Conditions for Streams in the Mid-Atlantic Highlands Region. Environmental Monitoring and Assessment 78(2):169-212.

Lazorchak, J.M., D.J. Klemm, and D.V. Peck (Editors), 1998. Environmental Monitoring and Assessment Program-Surface Waters. Field Operations and Methods for Measuring the Ecological Condition of Wadeable Streams, EPA/620/R-94/004F, U.S. Environmental Protection Agency, Washington, D.C.

Loch, D.D., J.L. West, and D.G. Perlmutter, 1996. The Effect of Trout Farm Effluent on the Taxa Richness of Benthic Macroinvertebrates. Aquaculture 147(1-2):37-55.

Manly, B.F.J., 1997. Randomization, Bootstrap and Monte Carlo Methods in Biology. Chapman and Hall, New York, New York.

McDonald, M., R. Blair, D. Bolgrien, B. Brown, J. Dlugosz, S. Hale, S. Hedtke, D. Heggem, L. Jackson, K. Jones, B. Levinson, R. Linthurst, J. Messer, A. Olsen, J. Paul, S. Paulsen, J. Stoddard, K. Summers, and G. Veith, 2004. The Environmental Protection Agency's Environmental Monitoring and Assessment Program. *In:* Environmental Monitoring. G. B. Wiersma (Editor). CRC Press LLC, New York, New York, pp. 649-668.

McNeil, W.J. and W.H. Ahnell, 1964. Success of Pink Salmon Spawning Relative to Size of Spawning Bed Materials. U.S. Fish and Wildlife Service Special Scientific Report, Fisheries 469, Washington, D.C.

Newcombe, C.P. and J.O.T. Jensen, 1996. Channel Suspended Sediment and Fisheries: A Synthesis for Quantitative Assessment of Risk and Impact. North American Journal of Fisheries Management 16(4):693-727.

Niemi, G.L. and M.E. McDonald, 2004. Application of Ecological Indicators. Annual Review of Ecology, Evolution, and Systematics 35:89-111.

Perry, J. and E. Vanderklein, 1996.Water Quality: Management of a Natural Resource. Blackwell Science, Cambridge, Massachusetts.

Peterson, N.P., A. Hendry, and T.P. Quinn, 1992. Assessment of Cumulative Effects on Salmonid Habitat: Some Suggested Parameters and Target Conditions. University of Washington, Center for Streamside Studies, TFW-F3-92-001, Seattle, Washington.

Plafkin, J.L, M.T. Barbour, K.D. Porter, S.K. Gross, and R.M. Hughes, 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers: Benthic Macroinvertebrates and Fish. U.S. EPA Report EPA/440/4-89/001, Washington, D.C.

Qian, S.S., R.S. King, and C.J. Richardson, 2003. Two Statistical Methods for the Detection of Environmental Thresholds. Ecological Modelling 166(1-2):87-97.

Rahlfs, V.W., 1997. Understanding and Evaluating Clinical Trials [Letter]. Journal of the American Academy of Dermatology 37(5):803-804.

Rausina, G.A., D.C.L. Wong, W. Raymon Arnold, E.R. Mancini, and A.E. Steen, 2002. Toxicity of Methyl Tert-Butyl Ether to Marine Organisms: Ambient Water Quality Criteria Calculation. Chemosphere 47(5):525-534.

Russo, R.C., 2002. Development of Marine Water Quality Criteria for the USA. Marine Pollution Bulletin 45(1-12):84-91.

Spooner, J., L. Wyatt, S.W. Coffey, S.L. Brichford, J.A. Arnold, M.D. Smolen, G.D. Jennings, and J.A. Gale, 1991. Nonpoint Sources. Research Journal of the Water Pollution Control Federation 63(4):527-536.

Stevens, D.L., Jr., 1997. Variable Density Grid-Based Sampling Designs for Continuous Spatial Populations. Environmetrics 8:167-195.

Stevens, D.L., Jr. and A.R. Olsen, 1999. Spatially Restricted Surveys Over Time for Aquatic Resources. Journal of Agricultural, Biological, and Environmental Statistics 4(4):415-428.

Suttle, K.B., M.E. Power, J.M. Levine, and C. McNeely, 2004. How Fine Sediment in Riverbeds Impairs Growth and Survival of Juvenile Salmonids. Ecological Applications 14(4): 969-974.

Tappel, P.D. and T.C. Bjornn, 1983. A New Method of Relating Size of Spawning Gravel to Salmonid Embryo Survival. North American Journal of Fisheries Management 3:123-135.

USEPA (U.S. Environmental Protection Agency), 1994. Water Quality Standards Handbook (Second Edition). EPA/823/B-94/005, U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA (U.S. Environmental Protection Agency), 2000a. Nutrient Criteria Technical Guidance Manual: Rivers and Streams. EPA-822-B-00-002, U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA (U.S. Environmental Protection Agency), 2000b. Mid-Atlantic Highlands State of the Streams. EPA/903/R-00/005, U.S. Environmental Protection Agency, Region 3, Philadelphia, Pennsylvania.

USEPA (U.S. Environmental Protection Agency), 2002. National Water Quality Inventory: 2000 Report. EPA-841-R-02-001, U.S. Environmental Protection Agency, Office of Water, Washington, D.C.

USEPA (U.S. Environmental Protection Agency), 2003a. Developing Water Quality Criteria for Suspended and Bedded Sediments (SABS). Potential Approaches: A U.S. EPA Science Advisory Board Consultation. U.S. Environmental Protection Agency, Office of Water, Draft Report. *Available at* http://www.epa.gov/waterscience/criteria/sediment/sab-discussion-paper.pdf. *Accessed in* August 2005.

USEPA (U.S. Environmental Protection Agency), 2003b. Setting Priorities to Strengthen the Foundation for Protecting and Restoring the Nation's Waters, Strategy for Water Quality Standards and Criteria. EPA-823-R-03-010, Office of Science and Technology, Office of Water, U.S. Environmental Protection Agency, Washington, D.C.

USEPA (U.S. Environmental Protection Agency), 2004. Environmental Monitoring and Assessment Program (EMAP): Mid-Atlantic Streams 1993-96 Data Sets. *Available at* http://www.epa.gov/emap/html/datal/surfwatr/data. *Accessed in* August 2005.

Venables, W.N. and B.D. Ripley, 1997. Modern Applied Statistics With S-Plus (Second Edition). Springer, New York, New York.

Waite, I.R., A.T. Herlihy, D.P. Larsen, and D.J. Klemm, 2000. Comparing Strengths of Geographic and Nongeographic Classifications of Stream Macroinvertebrates in the Mid-Atlantic Highlands, USA. Journal of the North American Benthological Society 19:429-441.

Yin, D., H. Jin, L. Yu, and S. Hu, 2003. Deriving Freshwater Quality Criteria for 2,4-Dichlorophenol for Protection of Aquatic Life in China. Environmental Pollution 122(2):217-222.

Young, M.K., W.A. Hubert, and T.A. Wesche, 1991. Selection of Measures of Substrate Composition to Estimate Survival to Emergence of Salmonids and to Detect Changes in Stream Substrates. North American Journal of Fisheries Management 11:339-346.

Zweig, L.D. and C.F. Rabeni, 2001. Biomonitoring for Deposited Sediment Using Benthic Invertebrates: A Test on 4 Missouri Streams. North American Benthological Society 20(4):643-657.

EXHIBIT 8

06/11/08

MEMORANDUM

From: William Swietlik, Chief
Ecological and Health Processes Branch
EPA/OW/HECD/OST

To: Robert Koroncai, Associate Director
Water Protection Division
EPA/Region 3

Subject: Development of Nutrient Endpoints for TMDLs in Pennsylvania

The Headquarters nutrient team and I have completed our review of the document entitled: *Development of Nutrient Endpoints for the Northern Piedmont Ecoregion of Pennsylvania*, prepared for Region 3 by Tetra Tech, Inc., dated November 20, 2007. It is our conclusion that the approach used in the document to derive the nutrient TMDL endpoints for use in implementing Pennsylvania's narrative standard is a scientifically defensible approach and is consistent with EPA guidance for deriving nutrient criteria.

The approach used in the document is an example of the multiple-lines-of-evidence (or weight-of-evidence) approach. The report examined different lines of evidence to derive nutrient numbers in three categories, and involved 17 different lines of evidence, constituting a very thorough analysis. The multiple-lines-of evidence approach is recommended by EPA in the following guidance.

- U.S. EPA. 2000b. *Nutrient Criteria Technical Guidance Manual: Rivers and Streams*. Office of Water, Office of Science and Technology, Washington, DC. EPA-822-B-00-002. In summary, this guidance states that a weight of evidence approaches that combines one or more of the three approaches; 1) Reference reaches, 2) Predictive relationships and, 3) Published threshold values; while considering downstream effects, will produce criteria of greater scientific validity.

- U.S. EPA. 2006. *Framework for Developing Suspended and Bedded Sediments (SABS) Water Quality Criteria.* Office of Water and Office of Research and Development, Washington, DC. EPA-822-R-06-001. This document recommends an integration and synthesis of multiple methods. This recommendation is based on a conclusion of the USEPA Science Advisory Board that "no single method would suffice complete criteria development in every situation and that multiple methods applied simultaneously (synthesized) may be more appropriate for criteria development."

In October, 2007 EPA HQ provide training to the Region II and III States on the weight-of-evidence methodology and how it can be applied to developing numeric nutrient values. It is good to see the Region benefited from our training and that you are now employing this approach.

EXHIBIT 9

UNIVERSITY OF DELAWARE

Department of Civil and Environmental Engineering
University of Delaware | Newark, DE 19716-3120
phone: 302-831-2442 | info@ce.udel.edu | fax: 302-831-3640

6/17/2008

John C. Hall, Esq.
Hall and Associates
1101 15th Street, NW
Washington, DC 20005

Re: Determining Appropriate Nutrient Reduction Requirements for Streams

Dear Mr. Hall:

This letter is in response to your recent inquiry regarding the establishment of nutrient standards to protect stream environments. I have devoted over 35 years of my professional career to the evaluation of nutrient impacts and other chemical interactions related to water quality and sediment criteria development. Nutrient standard development is a complex subject that is not amenable to simplified determinations. Stream environments, in particular, are subject to a variety of physical conditions that alter whether and how nutrients may stimulate excessive plant growth and significantly alter stream ecology. Any defensible analysis would have to account for these factors. This is a widely held view that is reflected in numerous articles in the peer reviewed literature and federal guidance documents.

You asked in particular for my views on a recent federal approach suggested for use in deriving nutrient standards for streams in Pennsylvania. As I understand it, the approach seeks to directly correlate total phosphorus (TP) "growing season" average concentrations with monitored invertebrate populations. It proposes to use a procedure known as "conditional probability" to identify an instream TP target that would protect invertebrate populations. You provided me with one such report that includes such an analysis. You asked, in general, whether the scientific literature would support the using such a direct correlation for development of a protective instream standard. The following provides a brief response

1. **Is it an accepted principle in the scientific literature that total phosphorus directly impacts invertebrate levels in streams?**

Answer: No. Nutrients do not act like toxic chemicals, e.g. copper. They do not have a direct impact on sensitive invertebrates. Therefore, it is not appropriate to directly compare nutrient levels in general and TP in particular to invertebrate responses. A scientifically defensible analysis must show how nutrients are affecting plant growth and then, how such plant growth is adversely impacting the ecology. It is well recognized in the literature that nutrients may cause an adverse impact to lake ecology if excessive algal productivity occurs. The relationships between nutrient loadings (*not* nutrient concentrations) and algal productivity in lakes can be analyzed using a variety of simple and more complex models. The mechanisms by which excessive algae cause problems include the effects of low dissolved oxygen, reduced transparency and loss of fish habitat. The analogous relationships for streams are less well understood and documented. To document invertebrate impairments in streams that are related to excessive nutrients, it is first necessary to determine the level of nutrients that is causing excessive plant growth in that stream and then to relate the consequences of excessive plant growth, e.g. low dissolved oxygen, to its effect on stream biota.

2. **Does a correlation between total phosphorus concentrations and invertebrate levels demonstrate that the nutrient was the cause of the changing invertebrate level?**

Answer: No. First, correlation does not demonstrate causation and therefore, unless more is known (i.e., the elevated phosphorus is also documented to cause excessive plant growth) such analysis gives a preliminary indication, at best. Use of a TP concentration indicator for stream environments, in particular, is subject to confounding factors. Elevated total phosphorus may occur due to high erosion rates that are inimical to sensitive invertebrate because they destroy the habitat – a common problem encountered in smaller streams. Simply plotting total phosphorus versus invertebrate population could easily be misleading and result in regulating the wrong water quality parameter and the wrong source of impairment. Second, the use of total phosphorus is not the best indicator of the form of phosphorus that could stimulate excessive fixed plant growth in streams. It is widely understood that dissolved phosphorus, not total phosphorus, is the form that is used by plants. Therefore, such an analysis cannot prove that phosphorus is causing the identified change in invertebrate populations.

3. **What is the accepted scientific approach to show whether or not nutrients could be causing an impact on invertebrates in a stream?**

Answer: Develop a model that accounts for the relevant physical, chemical and biological factors influencing how a stream responds to nutrient inputs. Then relate the effects of increasing plant growth to the response of organisms expected to exist in the habitat in question. Such an evaluation could utilize a correlation type of assessment but only if the relevant physical and chemical factors were accounted for in the evaluation and only for waters with closely similar ecology were being reviewed. One important factor to consider is stream orientation, shading and water transparency as it is widely understood that the amount of incident light influences the degree of primary production occurring, regardless of the amount of nutrients present. The physical habitat (cobble, sandy, or rocky bottom) will also influence the types of sensitive invertebrates that may be present at a location, independent of water quality. Finally stream flow itself is an important factor.

In closing I would like to address one further issue, the appropriate application of conditional probability methodologies for setting criteria and standards. The conditional probability approach suggested for use in Pennsylvania has, to my knowledge, never been used to derive a federal numeric water quality objective. The problem is whether the variable in question, total phosphorus concentration, is directly related to the effect that is being evaluated. Or to put it another way, is it clear that reducing the total phosphorus concentration would reduce the adverse effect, i.e. loss of benthic biota. This issue arose during the Science Advisory Board (SAB) review of the sediment criteria for which I was the technical director on behalf of EPA. We rejected this approach in favor of a method that directly links cause to effect.

One way to judge whether the conditional probability method is useful is to examine the probably plot itself. Note that most of the probability of an adverse effect ranges from 40% at the smallest TP concentration (5 ug/L) to 70% for approximately 100 ug/L.. Therefore large changes in concentration would be predicted to have only a small change in adverse effect probability: 70% to 40% for a twenty fold concentration change. It is likely that such a small change in adverse effect probability would not be detectable. Furthermore there are data used in the probability plot where *increasing* the total phosphorus causes a *decrease* in adverse effect probability. The reason for these is that there is only a weak relationship between TP concentration and adverse effects. Many other factors influence the benthic biota.

Therefore, while conditional probability can be considered as a possible tool for identifying streams for further investigation, it was not a sufficient basis to establish standards. In my opinion, such an approach cannot provide information on the pollutant level and form (e.g. dissolved or total phosphorus) that must be regulated to protect the environment.

One final comment: using this statistical procedure would be contrary to all the EPA criteria development that has preceded this effort. The scientific basis for the EPA water quality criteria is one of the landmark achievements of the agency. The methodology has been adopted almost universally. It has been reviewed many times. It is based squarely on the causal relationships between the chemical being regulated and the effects being protected. The proposed nutrient criteria would be a retreat from scientifically defensible criteria to simply an expedient solution for which little or no support exists. The claim that EPA employs the "best science" would not be true in this case. At the minimum I would strongly urge that an EPA Science Advisory Board or a National Academy of Science peer review be conducted to provide an independent evaluation of the proposed methodology.

Please let me know if I can provide any further information.

Sincerely,

Dominic M. Di Toro

Dominic M. Di Toro
Edward C. Davis Professor of Civil and Environmental Engineering
University of Delaware
Member of the National Academy of Engineering

EXHIBIT 10

TUFTS UNIVERSITY
School of Engineering

Professor and Louis Berger Chair in Computing and Engineering

July 8, 2008

John C. Hall, Esq.
Hall and Associates
1101 15th Street, NW
Washington, DC 20005

Dear Mr. Hall:

This letter is in response to your inquiry regarding the determination of appropriate nutrient reduction requirements for streams. Since 1974 when I first began studying the impact of phosphorus on Great Lakes water quality, I have devoted a great part of my scholarship to eutrophication. Although I initially focused on lakes and reservoirs, I shifted my focus to streams when I joined the faculty at the University of Colorado in 1986. It was there that I first recognized the significant and sometimes subtle differences between lakes and flowing systems like rivers and streams. Among other things, this resulted in my developing EPA-sponsored software that is expressly designed to simulate stream eutrophication (Chapra et al 2008).

Note that I have carefully read Prof. Di Toro's critique and wholeheartedly concur with all his conclusions. In particular, I strongly support his observations regarding the complexity of linking total P and invertebrate levels in streams. I also concur with his more general contention that the accepted scientific approach involves using models to quantify how streams respond to nutrient loadings. In the following, I will therefore limit my remarks to concerns that supplement his comments.

As I understand it, the proposed approach seeks to directly correlate total phosphorus (TP) concentrations with the health of invertebrate populations. Thus, phosphorus is treated as if it were a toxic substance that directly interferes with the viability and functioning of the biota.

Department of Civil and Environmental Engineering
223 Anderson Hall
Medford, Massachusetts 02155
617 627-3654
Fax: 617 627-3994
Email: steven.chapra@tufts.edu

This is such a scientifically indefensible representation of the connection between nutrients and ecosystem health that I believe that its adoption would represent a grave mistake. Beyond being vulnerable to legal challenge, I am much more concerned that its adoption would ultimately be ineffective. That is, it could lead to costly controls that would not protect our precious stream ecosystems.

An analogy with one of the earliest water-quality management problems is instructive in illustrating my concern. In the early 20th century, point discharges of untreated urban sewage resulted in low oxygen concentrations in many of our nation's rivers and estuaries. Because adequate oxygen levels are necessary for most forms of aquatic life, the stress on ecosystems was great. Aside from direct impacts on fish and macroinvertebrates, low oxygen also triggers a number of secondary effects such as the generation of noxious odors.

Water-quality engineers recognized early on that the root cause of oxygen depletion was the presence of organic carbon and reduced nitrogen compounds in the urban wastewater. In the simplest sense, these compounds were collectively called biochemical oxygen demand or BOD. Once introduced into the river, the BOD was broken down by bacteria. Because the bacteria consumed oxygen to affect the break down, the river's oxygen resources were depleted.

As depicted in Figure 1*a*, the cause-effect sequence for this problem is BOD loading → BOD concentration → dissolved oxygen concentration deficit[1] → biotic impact. Notice that along with BOD oxidation, the deficit is also determined by gas transfer or reaeration.

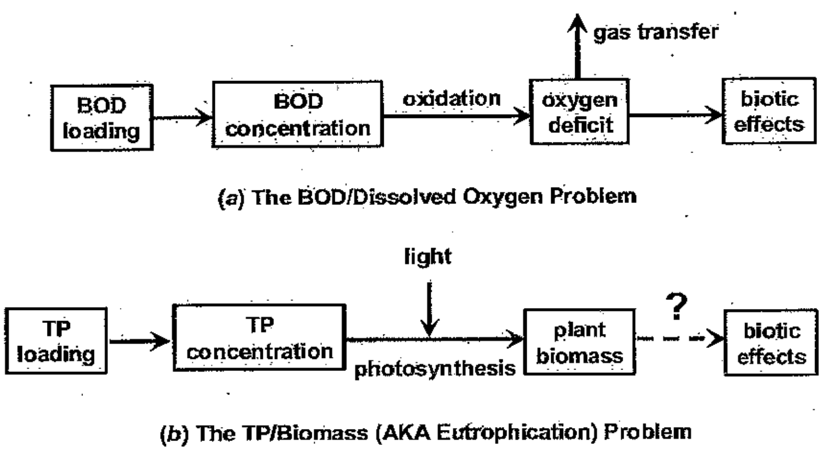


Figure 1 Analogy between two river water-quality problems: (a) dissolved oxygen and (b) eutrophication. The question mark on the link between plant biomass and biotic effects is meant to suggest the complexity of this connection.

Although this is a very simplified representation of the problem, the important point is that the direct cause of the ecosystem impairment is the oxygen deficit, not the BOD concentration. Hence, although we might ultimately alleviate the ecosystem stress by reducing the BOD loading, the actual BOD concentration would in fact not be directly correlated with the low oxygen levels.

[1] The deficit measures the difference between the saturation concentration and the actual DO level (Figure 2). A high deficit, therefore, indicates low oxygen and a high ecosystem stress.

This can be clearly seen by using the classic Streeter-Phelps model to develop a plot of BOD and oxygen in a river below a single point source of BOD. Although this model is simple, it accurately captures the interplay between BOD and oxygen in one-dimensional rivers[2].

As in Figure 2, the BOD is highest at the discharge point ($t = 0$) where the BOD loading is introduced. It then decreases downstream as the BOD is broken down via oxidation. The BOD oxidation in turn consumes oxygen which leads to a rapid decrease in oxygen concentration. However, as the level plunges, oxygen transfer across the air-water interface increases. Eventually the oxygen profile levels off when the oxidation loss is balanced by reaeration. Thereafter, oxygen levels begin to climb as reaeration becomes dominant and the stream recovers.

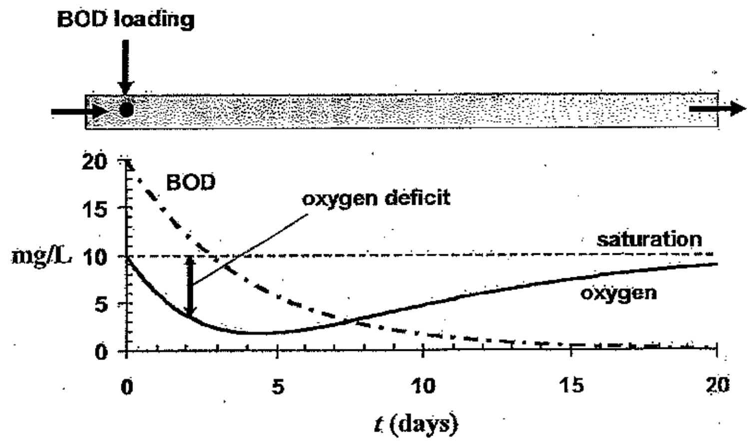


Figure 2 Simulation of BOD and dissolved oxygen versus travel time below a single point source into a one-dimensional river.

The most important feature of Figure 2 is that there is absolutely no spatial correlation between the in-stream BOD and oxygen levels. For example the highest BOD concentration and the lowest deficit both occur at the mixing zone ($t = 0$). Although BOD certainly causes oxygen depletion, it would therefore be ludicrous to specify an instream BOD concentration criterion in order to ensure an adequate oxygen level. Instead, the correct approach is to set an oxygen criterion that is directly connected with ecosystem health. A model can then be employed to link the oxygen concentration back to the BOD loading. This, of course, is how oxygen has been so effectively managed over the past century.

Beyond illustrating how river BOD and oxygen concentrations evolve and interact, employing a cause-effect model yields additional benefits. For example, rather than reducing the BOD load, the analysis suggests an alternative means to raise oxygen levels might involve enhancing gas transfer (e.g., oxygen diffusers, artificial waterfalls, etc.). Thus, the scientific approach reveals possible alternative remediation strategies that in certain cases might actually be more cost effective than source controls.

---

[2] In this context, one-dimensional means that changes only occur longitudinally and that the stream is well-mixed laterally (bank-to-bank) and vertically (with depth).

So how is this example relevant to the stream eutrophication problem? As illustrated in Figure 1*b*, a very simple representation of stream eutrophication is that nutrient loading (in this example, phosphorus[3]) results in increased stream nutrient concentration which in turn leads to increased plant biomass. Thus, as BOD is to oxygen deficit, phosphorus is to plant biomass. And just as it would be ludicrous to specify an instream BOD concentration criterion to solve the oxygen problem, it is equally misguided to specify an instream total phosphorus concentration criterion to solve the eutrophication problem.

Further, as described by Prof. Di Toro, the subsequent connection between increased biomass and biotic impacts is not as well understood as for oxygen. In fact, one such connection involves the impact of excessive plant growth on stream oxygen via both direct (photosynthetic gains and respiration losses) and indirect (when plants die they become BOD) pathways. However, the deleterious effects undoubtedly involve other non-oxygen related factors such as direct habitat impairment and shifts to undesirable plant species.

As with the previous oxygen example, beyond illustrating how river nutrients and biomass evolve and interact, the cause-effect model can point to alternative remediation approaches. For example, rather than reducing the nutrient load, Figure 1*b* suggests that an alternative means to reduce biomass might involve decreasing solar radiation. For example, this could be accomplished by planting riparian vegetation to create a canopy over the stream. As with oxygen, there might be cases where such alternatives could prove useful.

As a final technical note, it is important to distinguish between floating (i.e., phytoplankton) and attached (i.e., periphyton, filamentous algae and macrophytes) plants when dealing with stream eutrophication. For phytoplankton, which tend to dominate in deeper rivers, the direct analogy expressed in Figure 1 holds (see App. 1 for a detailed analysis). That is, it is absurd to set a TP criterion in order to manage river phytoplankton biomass.

For attached plants, the situation is much more complex. For systems dominated by macrophytes, nutrient management must consider whether the plants can draw nutrients from the sediments via their roots. For such situations, managing instream TP concentration would be counterproductive as photosynthesis would be effectively independent of water nutrients.

For filamentous algae (e.g., Cladophora) which draw nutrients directly from the water, there is certainly a closer connection. However, it is well known that (a) the photosynthesis rate of such organisms depends on their internal nutrient levels and (b) they only take up dissolved inorganic nutrients from the water. Hence, regulation based on water TP concentration would be ill-founded.

The same point can be made for periphyton but with a additional nuances. Whereas filamentous algae extend up into the water column, periphyton grow as biofilms on substrates such as bottom rocks. In such cases, their nutrient uptake can be influenced by transport limitations on the delivery of dissolved inorganic nutrients from the water into the biofilm. Such limitation would be dependent on stream hydraulics and hence related to factors such as stream velocity, depth, etc. Further, because they are bottom dwellers, their growth would obviously be highly dependent on the delivery of light to the stream bottom. Hence, along with nutrients, the clarity and depth of the overlying water would have to be considered in determining their

---

[3] Notice that in Figure 1, we have assumed that phosphorus is the limiting nutrient. It should be understood that this is an assumption and that another nutrient (e.g., nitrogen) or light might in fact be limiting.

biomass. Consequently, a simple, direct correlation with water TP concentration would seem to be overly simplistic.

In conclusion, I hope that the foregoing provides some indication of my great concern over the issue of stream eutrophication management. As a concerned environmentalist, as well as a lifelong fly fisherman, I truly support effective regulations to protect our nations great rivers and streams. However, as an environmental scientist and engineer, I also know that without a sound scientific basis, such regulations are likely to fail. It is in this spirit that I strongly support Prof. Di Toro's suggestion that at the minimum, an EPA Science Advisory Board or a National Academy of Science peer review be conducted to provide an independent evaluation of the proposed nutrient criteria.

Please let me know if I can provide any further information.

Sincerely,

Steven C Chapra

Steven C. Chapra, Ph.D.

APPENDIX 1.
Why TP Concentration Standards are Inappropriate for Managing Phytoplankton Biomass in Rivers

This appendix attempts to address the question of why anyone would ever suggest that a total phosphorus criterion would represent a sensible strategy for managing river and stream eutrophication dominated by phytoplankton. In brief, I believe that the idea of river total phosphorus criteria originates from the misguided notion that effective lake management approaches can be seamlessly (and thoughtlessly) transferred to rivers and streams.

In the late 1960's and early 1970's, several limnologists suggested that total phosphorus concentration could serve as an effective trophic state indicator. In particular, Richard Vollenweider posited that lakes with total phosphorus concentrations less than 10 μgP/L would tend to be oligotrophic whereas those with greater than 20 μgP/L would tend be eutrophic.

Although Vollenweider himself repeatedly stated that these were approximate guidelines and not hard thresholds, the values were adopted by many lake managers as quantitative goals for managing lake eutrophication. And in fact, the approach has been a useful component of nutrient remediation schemes for a number of important systems including the Laurentian Great Lakes.

So why might the approach work for lakes and not for streams? The answers to this question lies in fundamental differences between these two types of natural waters.

In effect, the viability of the Vollenweider approach is predicated on the functioning of the particular lakes he studied. In particular, the approach was developed for deep, stratified, phosphorus-limited, North-temperate, lakes with long residence times (i.e., greater than a year). In such lakes, Vollenweider (and others) assumed that the spring total phosphorus concentration was a prime controlling factor of plant production over the ensuing summer growing season.

For this assumption to strictly hold, once the lake stratifies in late spring, the epilimnion must essentially behave as a batch or closed system. Thus, plant growth over the ensuing summer is primarily dictated by the finite store of nutrient represented by the spring phosphorus concentration rather than by external loads. The average summer level of biomass is then determined by the recycle of this pool between inorganic and organic forms. Empirical support for the approach was provided by a number of empirical correlations. The chief examples of these were logarithmic plots suggesting strong correlations between summer average chlorophyll *a* concentrations and spring total phosphorus concentration.

A simple computation can be used to illustrate how such an approach breaks down in rivers and streams. First, total phosphorus can be divided into three components

$$TP = p_p + p_i + p_o \qquad (1)$$

where p_p = phytoplankton phosphorus (μgP/L), p_i = inorganic phosphorus (μgP/L), and p_o = non-phytoplankton organic phosphorus (μgP/L). If the chlorophyll a to phosphorus ratio is assumed to be 1 μgA/μgP, this means that p_p can be directly interpreted as a measure of phytoplankton biomass.

The river can be idealized as a steady-state, plug-flow system with a single point source of phosphorus (Figure 3). Further it is assumed that the river has uniform, steady flow and constant hydrogeometric properties (i.e., depth, width, etc.). For such cases, velocity will be constant and travel time and distance are linearly related (i.e., distance = velocity times travel time). Under these conditions, the following mass-balances can be written for each phosphorus component as

$$\frac{dp_p}{dt} = k_g \frac{p_i}{k_{sp} + p_i} p_p - k_r p_p - k_d p_p - k_s p_p \tag{2}$$

$$\frac{dp_i}{dt} = -k_g \frac{p_i}{k_{sp} + p_i} p_p + k_r p_p + k_h p_o \tag{3}$$

$$\frac{dp_o}{dt} = k_d p_p - k_h p_o \tag{4}$$

where t = travel time (d), k_g = maximum growth rate at constant light and temperature (/d), k_{sp} = phosphorus half-saturation constant (μgP/L), k_r = respiration/excretion rate (/d), k_d = death rate (/d), k_s = settling rate (/d), and k_h = hydrolysis rate (/d).

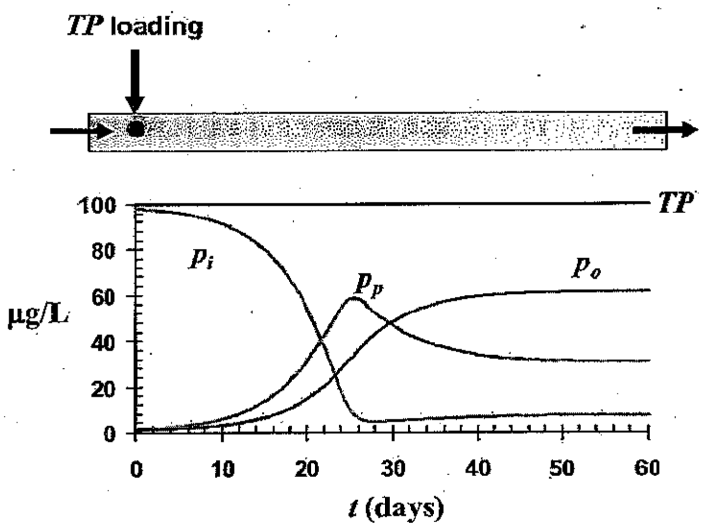


Figure 3 Simulation of phytoplankton, inorganic and organic phosphorus downstream from a point source.

Given reasonable values for the parameters and a set of initial conditions at the mixing point (Table 1), these equations can be integrated numerically to simulate how the various phosphorus species change as the water travels downstream. For the present example, the initial conditions are set so that the river has a high level of available, inorganic nutrient at the mixing point as would be the case for a high phosphorus discharge into an effluent-dominated river. In addition, the phytoplankton settling velocity is set to zero.

Table 1 Parameters and initial conditions used to simulate phytoplankton and phosphorus concentrations below a single point source to a one-dimensional river.

Parameter	Value	Units
k_g	0.5	/d
k_{sp}	5	µgP/L
k_r	0.2	/d
k_d	0.1	/d
k_s	0	/d
k_h	0.05	/d
Initial conditions:		
p_p	1	µgP/L
p_i	98	µgP/L
p_o	1	µgP/L

As displayed in Figure 3, because the inorganic P concentration is well above the half-saturation constant, the phytoplankton initially grow rapidly as the inorganic phosphorus is efficiently converted to phytoplankton biomass. Growth continues until the inorganic phosphorus level approaches the half saturation constant whereupon a peak is reached. At this point, growth has become sufficiently limited that it is exactly balanced by the respiration and death losses. Thereafter, the phytoplankton levels decline until the solution approaches a stable steady state. This asymptote represents the point at which phytoplankton growth exactly balances phosphorus recycle.

Note that because of the assumption of zero settling, the total P concentration is constant. This allows the component concentrations at the stable steady state to be computed exactly as

$$p_i = \frac{k_r + k_d}{k_g - (k_r + k_d)} k_{sp} \tag{5}$$

$$p_o = \left(1 - \frac{k_h}{k_d + k_h}\right)(TP - p_i) \tag{6}$$

$$p_p = \frac{k_h}{k_d + k_h}(TP - p_i) \tag{7}$$

Thus, we see that the ultimate inorganic phosphorus concentration is equal to the half saturation constant multiplied by the ratio of the phytoplankton loss rates ($k_r + k_d$) to the maximum net phytoplankton growth rate ($k_g - k_r - k_d$). The organic P and phytoplankton P concentrations are then dictated by the product of the total organic P (i.e., organic P and phytoplankton P) and a dimensionless number quantifying the relative values of the hydrolysis and death rates.

Although this is a very simple model, it dramatically illustrates why specifying a phosphorus concentration standard for rivers is ill-founded. Notice that until the asymptote is approached, there is no direct correlation between phytoplankton biomass and the total phosphorus concentration (as well as with any of the individual phosphorus species).

Just as is the case for BOD and oxygen, although phosphorus certainly causes increased phytoplankton biomass, there is absolutely no direct spatial correlation between in-stream TP and biomass. Hence, whereas a phosphorus standard makes some sense for a long residence time lake, it falls apart for a plug-flow system like a river.