April 28, 2010

Water Docket
U.S. Environmental Protection Agency
Mailcode: 2822T
1200 Pennsylvania Avenue, NW.
Washington, DC 20460
Via Electronic Mail: ow-docket@epa.gov

Attention: Docket ID No. EPA–HQ–OW–2009–0596

Dear Sir or Madam:

The National Association of Clean Water Agencies (NACWA) appreciates the opportunity to comment on *Water Quality Standards for the State of Florida's Lakes and Flowing Rivers Notice of Proposed Rulemaking*, published by the Environmental Protection Agency (EPA) in the *Federal Register* on January 26, 2010 (75 *Fed. Reg.* 4174). NACWA represents the interests of the nation's publicly owned wastewater treatment agencies, including 12 utilities in Florida. Beyond the obvious impacts on NACWA's members in the state, the proposal introduces new concepts that will likely impact clean water agencies nationwide as states continue to develop and implement numeric nutrient water quality criteria. Accordingly, NACWA has conducted a comprehensive review of the proposal and offers the comments and recommendations below. New approaches discussed in the proposal show promise in aiding implementation of nutrient controls, but concerns remain over whether they will work in real-world applications. In addition, NACWA believes that the situation in Florida exemplifies the Association's over-arching concerns with EPA's rigid approach to setting nutrient criteria. NACWA's members feel strongly that new approaches are needed that recognize the unique aspects of nutrients in the water environment.

The developments in Florida are particularly troubling given the scope and level of sophistication of the program that is already in place there. Florida is an authorized state under the Clean Water Act (CWA) and has adopted narrative water quality standards for nutrients that have been approved by EPA. Though perhaps not at a pace that satisfied all stakeholders, the Florida Department of Environmental Protection (FDEP) had made tremendous progress in developing and implementing total maximum daily loads (TMDLs), with meaningful controls for point and nonpoint sources, based on its narrative standards and using site-specific numeric targets. FDEP had also initiated an effort to develop numeric nutrient criteria for the

National Association of
Clean Water Agencies
1816 Jefferson Place, NW
Washington DC 20036-2505

p 202.833.2672 f 202.833.4657
www.nacwa.org · info@nacwa.org

state. Florida is a leader on this issue and, unfortunately, a legal challenge of the state's program has led to an aggressive schedule and use of criteria development procedures that largely ignore in-stream conditions and biological responses.

The pace of nutrient goal development has been generally slow nationwide, as states struggle to develop scientifically based approaches that will meet with EPA's approval. It will not take long for the legal strategy employed in Florida to spread to other states, and Florida will likely not be the last state where EPA is forced to establish nutrient criteria. How EPA has approached this issue in Florida, therefore, is relevant for the entire nation.

New Approaches Needed for Addressing Nutrients

Developing meaningful numeric nutrient criteria, particularly for flowing streams, is complex because of the many factors involved and cannot be accelerated to meet the demands of court-ordered deadlines without sacrificing the careful analysis required to accurately characterize specific waterbody responses to nutrients. Doing so only leads to criteria that may not accurately reflect the actual biological conditions of the waterbodies they are designed to protect. In the case of Florida, EPA has taken a statistical approach to developing nutrient criteria that assumed all of Florida's flowing waters could be grouped into four broad categories and that the behavior of all waterbodies in those categories will be the same. In order to establish numeric criteria for nitrogen and phosphorus, EPA dismisses the fact that it was unable to establish a cause and effect relationship between nutrient levels and in-stream impacts. This lack of a simple cause and effect relationship between nutrient concentrations and water quality illustrates the variability and complexity in individual waterbody responses to nutrient loadings.

Even where court-ordered deadlines are not forcing an acceleration of the scientific process, state regulators are having difficulty developing meaningful numeric criteria for nitrogen and phosphorus that are realistically linked with in-stream impacts. Where criteria are developed, many of the numeric values are so low that NACWA's members are concerned that if applied as end-of-pipe effluent limits, they would be difficult, if not impossible to meet with existing treatment technology. Numeric nutrient criteria that translate to effluent limits so low that wastewater utilities would need to treat their effluent with reverse osmosis, go well beyond what is realistically considered to be the "limit of technology" for nutrient removal and may harm the environment through excessive use of energy, chemicals, and emission of greenhouse gases. NACWA believes additional flexibility is needed for states to develop nutrient criteria that better reflect actual in-stream conditions and watershed nutrient loadings to best target the source of the problem and develop sustainable management plans to protect water quality.

Rapid development of numeric criteria for nitrogen and phosphorus that are close enough and providing implementation flexibility to compensate for any shortcomings in the criteria is not a sustainable approach. Some of the concepts EPA proposed with the Florida criteria, including the restoration standard, have promise for improving implementation. However, NACWA believes there may be alternative, more sustainable approaches that accomplish significant nutrient reduction while also providing greater flexibility and having less impact on a community's economic viability. The criteria development process should include a thorough consideration of implementation, including guidance on effluent discharger permitting, and states must have the flexibility to explore the use of criteria that don't simply include rigid numbers for total nitrogen (TN) and total phosphorus (TP).

Four Principles for the Development of Valid Nutrient Controls

NACWA believes the following principles are essential for the development of valid and meaningful nutrient controls and that ultimately a new, national paradigm for addressing nutrients is needed:

1. Water quality goals must be technically and scientifically defensible and technically achievable;
2. Water quality goals must be based on a demonstrated cause and effect relationship between stressor and response variables and should not be implemented until a biological impact has first been demonstrated;
3. A watershed approach using adaptive management is essential to meeting long-term water quality goals; and
4. Water quality goals should result in a net environmental benefit and consider cost and ancillary benefits in relation to effectiveness and sustainability.

Specific Recommendations for a Scientifically Valid Program

NACWA believes the approach taken in Florida by EPA side-steps these core principles in order to ensure the criteria fit EPA's established mold for developing water quality criteria. NACWA offers the following recommendations for moving forward with a more reasonable and scientifically valid program for addressing nutrient enrichment in Florida and the rest of the nation:

1. If EPA continues to pursue the federal criteria for Florida as proposed, the Agency should draft and submit for public comment implementation guidance before the criteria are promulgated;
2. Nationally, EPA must change its existing framework that necessitates the establishment of numeric criteria for TN and TP, which may not be appropriate for all waters;
3. EPA should better recognize and approve new and innovative approaches for expressing nutrient water quality goals currently being considered by some states;
4. Future water quality criteria development work must be supported by information on biological response to ensure criteria will actually protect designated uses;
5. Additional flexibility, both in criteria development and in permitting (e.g., averaging periods), to account for the unique behavior of nutrients must be provided;
6. Ultimately, EPA, the states and remaining stakeholders must work to establish a new, more holistic approach to addressing water quality problems that includes meaningful and equitable controls and implementation requirements for all nutrient sources.

Building on these principles and recommendations, NACWA believes a new dialogue must begin to outline more reasonable approaches to nutrient control.

NACWA to Convene Summit on Nutrients, Initiate Dialogue with EPA, States

NACWA's members across the country have been engaged in the development of water quality goals for years and have gained valuable knowledge and experience related to effective implementation approaches for wastewater treatment plants. NACWA believes that a new approach or suite of approaches is needed at the national level to address nutrients and that simply pointing the finger at EPA is not a sufficient response. In line with this, NACWA is convening a Summit of its membership to outline potential solutions. NACWA will

convene its Nutrient Summit this summer and stakeholders from EPA and the states will be invited to share their perspectives related to criteria development and implementation. NACWA will generate a white paper based on the Summit for use in initiating further dialogue with EPA and the states on developing a comprehensive, integrated, scientifically sound, and cost effective approach to improving water quality with respect to nutrients.

The Summit and resulting report will acknowledge the importance of addressing all sources of nutrients, but focus on near-term solutions that recognize that current CWA authorities to address the array of nutrient sources are limited. NACWA believes that any long-term solution to nutrient impacts must result in equitable and accountable controls for all sources, proportionate to their contribution. Little or no progress will be made unless all nonpoint source categories are part of the solution. NACWA is working on a parallel track to explore legislative solutions that support a watershed-based approach to water quality improvement that would accomplish this long-term goal.

Over the past several years numerous approaches have been suggested by environmental activist groups and the states in an effort to accelerate progress on nutrient control. NACWA understands that there are concerns about the rate of progress on controlling nutrients and believes that there must be a new and proactive dialogue among the key stakeholders to find solutions that will actually make sense, rather than simply reverting to blanket, one-size-fits-all approaches.

EPA's Proposed Criteria Will Not Ensure Florida Waters Achieve CWA Goals

Despite Florida's robust and functioning water quality based program, including EPA-approved narrative nutrient criteria implemented via the TMDL program, EPA issued a determination on January 14, 2009 (EPA 2009) that numeric nutrient criteria were needed to meet the goals of the CWA. EPA's authority to issue water quality standards in an authorized state is found in section 303(c)(4)(B) of the CWA, which authorizes EPA to issue regulations "setting forth a revised or new water quality standard" in any case where the Administrator of EPA determines that new standards are needed to meet the requirements of the Act. The criteria that EPA promulgates will become part of Florida's water quality standards sixty days after the final criteria are issued. NACWA believes that EPA's action is not warranted given the state's existing program and that EPA's action will in fact further complicate efforts in Florida to address nutrient over-enrichment.

EPA's January 2009 letter indicates that the Agency had determined that numeric nutrient criteria "are necessary to facilitate and expedite the identification of all nutrient impaired waters in Florida; thereby providing necessary protection for the State's designated uses, as required by the CWA." EPA indicated that the reason for its determination was not Florida's failure to develop numeric nutrient criteria, but rather the procedure Florida had been following. Specifically, EPA was concerned that Florida's reliance on a case-by-case interpretation of its narrative nutrient criterion through its Impaired Waters Rule was taking too long to implement. In fact, when EPA made its determination, FDEP was already in the process of working with stakeholders in the state to develop numeric nutrient criteria. The speed with which Florida was making progress was the driving force behind this action, not the water quality needs of Florida.

In its determination, EPA makes it clear that there is little flexibility in developing water quality criteria for nutrients, insisting that numeric values for TN and TP are essential components. NACWA believes that EPA's continued insistence that only state-wide numeric nutrient criteria for TN and TP can meet the requirement in

the CWA to protect designated uses of waterbodies is not supported by the science and is leading to nutrient requirements that will likely be unachievable and result in limited environmental benefit. EPA cites the need to address excessive plant growth as a driver for federal promulgation of numeric nutrient criteria for Florida. However, the way EPA has structured the proposal, stringent nutrient requirements will be imposed even when plant growth is not found to be excessive. EPA's proposal applies all criteria components independently and impairment occurs if TN or TP concentrations are exceeded, regardless of plant growth conditions.

While NACWA plans to discuss the issue of nutrient criteria development in more detail in its upcoming Summit, the scientific literature and EPA's own Science Advisory Board (SAB) continue to indicate that numeric nutrient criteria, as currently dictated by EPA, are not the only approach to protect water quality. The SAB's Ecological Processes and Effects Committee (EPEC), in its recent draft report from January 2010 draft report, stated:

> [t]he way in which EPA used results from mechanistic models to develop nutrient load reduction goals for the Gulf of Mexico (Mississippi River/Gulf of Mexico Watershed Nutrient Task force, 2008), and the way in which it is currently using mechanistic models for nutrient and sediment TMDLs for Chesapeake Bay, does not involve development or use of numeric nutrient criteria. The reason is that these mechanistic models (Scavia et al., 2004; Cerco and Noel, 2004) are load-response models, not empirical stressor-response models, and hence they obviate the need for numeric nutrient criteria because they directly link nutrient loads to response variables that represent water quality impairments (e.g., dissolved oxygen, chlorophyll, water clarity and acreage of submerged aquatic vegetation). (SAB 2010, at 5)

Florida's current narrative criterion for nutrients states that: "In no case shall nutrient concentrations of a body of water be altered as to cause an imbalance in natural populations of aquatic flora or fauna." Florida currently implements this criterion on a site-specific basis, identifying the level of nutrients that would cause an imbalance and then translating that level into site-specific numeric targets. FDEP evaluates site-specific data on the trophic state, chlorophyll *a* values, and other water quality conditions such as algal blooms, excessive macrophyte growth, and the state of submerged aquatic vegetation to determine if there are any impairments of beneficial uses. Rather than relying on simple numeric values, Florida's regulations include both thresholds for nutrient impairment based on response variables and a confirmatory step using a site-specific bioassessment.

The January 2010 SAB Report praises Florida's approach:

> We agree with the statement in the Florida Department of Environmental Protection's letter of September 4, 2009 (letter from Daryll Joyner, Florida Department of Environmental Protection to Thomas Armitage, Designated Federal Officer, EPA Science Advisory Board Staff Office) indicating that the "most scientifically defensible strategy for managing nutrients within the range of uncertainty is to verify a biological response prior to taking a management action." This risk/performance-based approach to setting nutrient criteria is evident not only in Florida's program, but also in those developed by California and Maine (Florida Department of Environmental Protection, 2009; Maine Department of Department of [sic] Environmental Protection, 2009; McLaughlin and Sutula, 2007). Those risk-based linkages are not addressed in either the Guidance or EPA's Nutrient Criteria Technical Guidance documents for Rivers (2000), Lakes/Reservoirs (2000), and Estuaries (2001). (SAB 2010, at 6)

EPA cannot show that its proposed numeric criteria are better suited to meet the goals of the CWA than the current approach employed by Florida because it cannot demonstrate that its criteria will ensure that the designated uses are met. In fact, because EPA has not identified the threshold level of nutrients required by the various waterbodies in Florida, EPA cannot demonstrate that its proposed nutrient standards will even have a beneficial effect on designated uses. Unlike standards for toxic chemicals, it cannot be assumed that increasingly stringent nutrient criteria will produce water quality improvements. There is a threshold at which more stringent nutrient criteria will actually harm water quality because nutrients are essential to a healthy aquatic ecosystem.

Several of the proposed Florida criteria values are not based on any documented biological impairment threshold. The range of conditions occurring in non-impaired waters, purportedly reflected by EPA's reference stream approach, does not demonstrate what the impairment threshold is. EPA's analysis fails to show that higher levels of nutrients, above the established reference condition criteria, would actually impair uses. In addition, EPA is proposing to regulate both TN and TP regardless of actual need. Available data confirms that lake impairments are generally caused by phosphorus while excessive nitrogen loads cause estuary impairments. EPA disregards this information and seeks to regulate both in all cases.

In order to develop scientifically defensible nutrient criteria for Florida, the threshold level of nutrients needed to support the designated use of each waterbody must be identified using site-specific data. EPA's criteria, developed using the reference condition approach, provide no connection to the actual needs of Florida's waterbodies and ignore the biological conditions of those waters necessary to support beneficial uses. EPA's criteria, therefore, will not help Florida waters meet CWA goals and will instead complicate implementation of Florida's robust existing program.

Florida Criteria Rely on Flawed Reference Condition Approach

Water quality experts, the SAB, and EPA's own scientists acknowledge that establishing a causal relationship between TN and TP levels and in-stream biological response variables is essential to developing meaningful nutrient criteria. In its EPEC draft report from January 2010 on EPA's use of statistical models for criteria derivation, the SAB noted that "understanding the causative link between nutrient levels and impairment is necessary in order to assure that managing for particular nutrient levels will lead to desired outcomes" (SAB 2010, at 4).

The SAB elaborated stating that:

> Without a mechanistic understanding and a clear causative link between nutrient levels and impairment, there is no assurance that managing for particular nutrient levels will lead to the desired outcome. There are numerous empirical examples where a given nutrient level is associated with a wide range of response values due to the influence of habitat, light levels, grazer populations and other factors. If the numeric criteria are not based upon well-established causative relationships, the scientific basis of the water quality standards will be seriously undermined. (SAB 2010, at 4)

While the guidance that the SAB was commenting on focuses on stressor-response models, the concerns expressed by the SAB regarding the need to establish cause and effect apply equally to the reference criteria

approach to establishing water quality standards. As noted above, EPA was not able to find "sufficient scientific support for" the use of stressor-response relationships in Florida waters for developing numeric nutrient criteria. EPA instead defaults to its reference condition/percentile-based approach to criteria derivation, which it used to develop national criteria guidance in the early 2000s.

Since the late 1990s when it was first proposed for use in criteria derivation, NACWA has been a vocal critic of EPA's reference condition approach. As indicated in an independent review (Appendix 1) commissioned by NACWA for its comments on EPA's nutrient criteria technical guidance for river and streams, "…evaluation based on frequency distributions of primary variables from a variety of streams even within some defined ecoregion simply does not contain information about the causal relationships between those variables and the potential for impairment of biological integrity in any given stream or reach" (HydroQual 2000, at 3). The review also stated that "it is this philosophical departure from effects-based criteria setting that is most disagreeable and which creates a fundamental problem of credibility for the nutrient criteria setting process" (HydroQual 2000, at 1). Statistical percentiles of a data set only represent the properties of a data distribution. Without groundtruthing and further analysis of associated effects on beneficial uses, it is not possible to ascertain whether the given percentile value reported as a criteria will address an impairment. This was confirmed most recently in the SAB's report on EPA's use of statistical tools to derive nutrient criteria.

EPA notes in the preamble to the Florida proposal that "[t]he reference condition approach relies on the identification of reference waters that exhibit minimal impacts from anthropogenic disturbance and are known to support designated uses. The thresholds of nutrient concentrations where designated uses are in attainment are calculated from a distribution of the available associated measurements of ambient nutrient concentrations at these reference condition sites." But FDEP has already raised concerns about EPA's selection of reference waters for the proposal. In its efforts to propose numeric nutrient criteria, Florida had contemplated using a reference condition approach, but would have excluded more waters from its list of reference sites based on the influence of anthropogenic sources of nutrients.

Florida also wanted the criteria set at a higher percentile (90th) of the statistical distribution rather than the 75th percentile used by EPA. As an alternative, Florida was considering requiring biological validation or verification before declaring impairment – which would have provided some reassurance that the statistically derived criteria were not unnecessarily restrictive. Other states have proposed similar approaches, numeric TN and TP values but with required biological verification, but EPA has so far not expressed support for such approaches.

EPA's selected reference waters and reliance on the 75th percentile will, by FDEP's estimates, result in 35 percent of Florida's most pristine surface waters failing EPA's proposed in-stream criteria. Peer review comments on EPA's *Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs* included disagreement about using the reference condition approach including using percentile values as arbitrary and lacking scientific rationale for criteria development. EPA's response at the time was that science cannot address all aspects of environmental management decision making. While NACWA recognizes that the science in some areas of environmental protection are not sufficiently advanced to provide adequate information on how or to what level pollutants should be controlled to protect designated uses, this is not the case with nutrients. There is sufficient science available to develop cause and effect relationships related to nutrients – the Chesapeake Bay nutrient TMDL is a

good example. Although the effort may be complex and difficult it is possible to model relationships rather than defaulting to over-simplified criteria like those proposed by EPA.

EPA's own guidance indicates that Agency scientists viewed reference condition-based criteria as "starting points" stating the "values for both causal (total nitrogen, total phosphorus) and biological and physical response (chlorophyll *a*, turbidity) variables represent a set of starting points for States and Tribes to use in establishing their own criteria in standards to protect uses" (EPA 2000, at 1). This statement was supported by the following additional disclaimer statements: "EPA expects that, in most cases, it will be necessary for States and Tribes to identify with greater precision the nutrient levels that protect aquatic life and recreational uses" (EPA 2000, at 5), and "[f]or example more sensitive uses may require more stringent values as criteria to ensure adequate protection. On the other hand, overly stringent levels of protection against the adverse effects of cultural eutrophication may actually fall below levels that represent the natural load of nutrients for certain waterbodies. In cases such as these, the level of nutrients specified may not be sufficient to support a productive fishery" (EPA 2000, at 5).

If EPA's reference condition-based criteria are truly intended to be 'starting-points', then the implications for Florida, where criteria could be used as end-of-pipe limits, are even more significant.

Proposed Reference Condition Criteria Are Not Statistically Defensible

Though NACWA believes the reference condition approach should not be used to develop criteria, NACWA worked with utilities in Florida to evaluate EPA's application of the reference condition approach methodology in the Florida proposal (See Appendix 2). Based on a review of the information used to develop EPA's proposal, NACWA believes the criteria suffer from a lack of sufficient underlying data. The instream protective criteria for streams proposed by the EPA rely upon insufficient data necessary to obtain a reliable estimate of a protective threshold. Fully 70 percent of the stations utilized by the EPA have data for only one year and another 15 percent have data for only two years during the 19-year period covered by the data. As a result of having such limited time coverage, the temporal (year-to-year) variability cannot be reliably measured, and therefore this critical component is missing from EPA's threshold calculation. Instead, the EPA simply ignores the year-to-year variability by utilizing stations aggregated over years, which, for 70 percent of the stations, is a single observation in a single year. This does not produce statistically defensible criteria that can be used to judge whether an annual site mean is or is not out of compliance.

The EPA also proposed an alternative methodology developed by the FDEP. It utilizes data collapsed to the waterbody-year level, averaged over stations and sampling events. With sufficient data, a statistical method that considers the spatial (station-to-station) and temporal (year-to-year) variability could be used to develop a reliable criterion. However, using aggregated data that ignores the variability inherent in the data will not reliably measure nutrient levels needed to regulate water quality.

Regardless of which set of standards are adopted – EPA or FDEP – the chance of declaring a site out of compliance when it is truly in compliance is a critical attribute of any test criteria. The false positive rates for water quality compliance criteria are important in evaluating the likelihood of undertaking extensive, expensive, but ultimately unnecessary, remedial actions. The EPA instream criteria are based upon the 75th percentile, which generates unacceptable false positive levels exceeding 25%, even assuming the criteria are calculated reliably, which, as explained above, they are not. The actual false positive rates may be much higher

due to the insufficient data, and improper data aggregation, employed by the EPA. This means that sites that are fully compliant will have at least a 25% chance of being declared out of compliance. Although the FDEP use of the 90th percentile will in theory result in more reasonable false positive rates, the criteria still must be based on sufficient data to be reliable.

Downstream Protection Methodology Needs Formal Review

For the first time, EPA is proposing an approach to modify instream water quality criteria for the protection of downstream uses. NACWA understands that EPA has tabled its decision on the downstream protective values relating to estuaries until next year, but given the potential implications for the nation, NACWA feels it is important to fully address the concept of developing downstream protective criteria values in these comments.

NACWA believes that what has been proposed was not sufficiently evaluated and oversimplifies the complex watershed dynamics that determine how nutrients are actually processed and delivered to downstream waters. NACWA has identified significant issues with the assumptions EPA has made in its approach to estimating downstream protective values and believes that the entire approach should be put on hold until an in-depth analysis of the issue can be conducted. NACWA is concerned that this approach, as contemplated in the proposal, would supersede any watershed-specific process that might better be used to develop appropriate goals based on site-specific information and mechanistic modeling and create a conflict between Section 303(c) and 303(d) of the CWA. Downstream protection is already achieved through the existing TMDL program – loadings from upstream discharges to impaired downstream waters are addressed through wasteload allocations.

For proposed water quality standards, a significant unresolved question for many states has been, how far downstream should beneficial uses be protected? The EPA proposal for Florida is a process that is essentially a total maximum annual load calculation on lakes and reservoirs to ensure that the criteria set for inflowing streams do not impair waters downstream. This method of determining downstream impacts is combining two distinct ideas, water quality standards and TMDLs. In the past these have been completely separate. Water quality standards have been based purely on the ability to meet the designated uses of the particular waterbody. EPA in effect, has set up a TMDL analysis standard for the state of Florida in the proposed rule. EPA has consistently taken the position that: "adoption of criteria that represent 'minimally impacted' conditions and that are sufficiently protective of near field downstream effects should ultimately achieve a far field benefit" (EPA, 2001, at 17). If further reductions are needed to protect downstream waters, those reductions should be based on a complete loading analysis with an allocation of nutrient loadings via the TMDL process.

NACWA has concerns about the validity (representativeness) of the underlying assumptions used in both the lake and estuary downstream protection value methodologies. In the proposal, EPA is seeking feedback on the approach and by referencing the limitations of the SPARROW model, EPA acknowledges that other approaches and models may also provide defensible estimates of protective loads. All of these considerations need to be reviewed before this downstream protection overlay is used in a regulatory or criteria-setting process.

As part of EPA's proposed methodology, TMDL calculations that have, or will be performed can be used to set the target criteria, which EPA refers to as a site-specific alternative criteria. It is not clear how the proposed downstream protective values would be used as the criteria for waterbodies where TMDL calculations have been

performed, or will be performed in the future. The uncertainty associated with changing water quality criteria associated with unknown future TMDL calculations is an additional concern to NACWA.

Finally, another unresolved issue relating to EPA's proposed approach for protecting downstream uses is its potential impact on upstream state standards. If EPA intends to apply the federal criteria in upstream states, it must fully engage those states in its rulemaking process.

EPA's Proposal Does Not Meet the Requirements of the Regulatory Flexibility Act and Unfunded Mandates Reform Act

In the proposal, EPA asserts that it is not required to determine the impacts on small businesses and small governments under the Regulatory Flexibility Act, or to consider and adopt the least costly, most cost-effective, or least burdensome alternative under the Unfunded Mandates Reform Act. EPA states that, when promulgated, federal water quality standards for nutrients in Florida's lakes and flowing waters, will have no regulatory effect on entities in the state of Florida. This may be the case for EPA's typical criteria guidance document development process, but EPA's proposed criteria for Florida are not guidance and will become standards in the state within sixty days of being published.

EPA's federally promulgated water quality standards will be legally applicable to small entities (as well as others) that discharge nutrients into Florida lakes and flowing waters. EPA claims that the "proposed rule does not regulate or affect any entity and, therefore, is not subject to the requirements of" the Unfunded Mandates Reform Act since the criteria will not be applied directly as a discharge limit for a permittee. However, where a discharge has the reasonable potential to cause, or contribute to an excursion of numeric or narrative water quality criteria, the permitting authority must develop permit limits as necessary to meet water quality standards. In waterbodies where a waste load allocation has not been prepared by a state and approved by EPA under 40 CFR 130.7, water quality-based effluent limits must ensure that the "level of water quality to be achieved by limits on point sources established under this paragraph is derived from, and complies with all applicable water quality standards" (40 C.F.R. 122.44(d)(1)(vii)(A)). Absent a TMDL, EPA has taken the position in the past that water quality based effluent limitations for discharges into impaired waters cannot exceed an applicable water quality criterion meaning the criteria will be applied as end-of-pipe limits:

> On its face, EPA believes that a WQBEL [water quality based effluent limit] allowing discharges into a waterbody already exceeding such criteria would not ensure that the water quality achieved by point sources was either "derived from" or "complies with" applicable water quality standards. EPA also believes that such a permitting approach would be fundamentally at odds with the water quality-based permitting requirement contained in section 301(b)(1)(C) of the CWA, since such an approach would allow point sources to contribute to the excursion above water quality standards in the waterbody. (EPA 1995, at 556)

This issue of applicability of a water quality criterion as an end-of-pipe limit is particularly significant where ambient instream concentrations exceed numeric nutrient criteria and in the arid west where a single wastewater treatment plant could comprise most or all of a waterbody's flow.

Though EPA downplayed the applicability of the Regulatory Flexibility Act and the Unfunded Mandates Reform Act to its actions, EPA did prepare an analysis of the costs of meeting the proposed federal standards as well as a "RFA/SBREFA Screening Analysis for Water Quality Standards for Nutrients in Lakes and Flowing Waters in Florida." As discussed below, EPA's cost estimate makes significant assumptions that result in a vast understatement of the cost impacts to dischargers in Florida. EPA's RFA/SBREFA Screening Analysis relies on this inadequate cost estimate to conclude that EPA's proposal will not result in a significant impact on a substantial number of small entities.

Utilities Face Significant Costs to Meet Nutrient Reduction Requirements

Given that EPA's proposed Florida nutrient criteria will become state standards and be used to impose effluent limits for dischargers, estimating the cost of EPA's proposal is both critical and complicated. EPA's cost estimate is detailed in a report entitled "Preliminary Estimate of Potential Compliance Costs and Benefits Associated with EPA's Proposed Numeric Nutrient Criteria for Florida" dated January, 2010. In this study, EPA noted that Florida already has in place narrative nutrient criteria, and had published July 2009 draft changes to its nutrient water quality standards to put in place numeric limits for phosphorous and nitrogen. EPA assumed that these standards were the baseline condition for its cost estimate, and that the state of Florida, in the absence of the federal proposed rule, would have finalized what it had proposed for state nutrient standards – this is a major assumption that has a significant impact on EPA's cost estimate. Dischargers in Florida had expressed significant concerns with FDEP's earlier proposal and it is not clear what changes may have been made to that proposal before being finalized.

The EPA cost report only identified 6 dischargers (all in the South region) and estimated that the total incremental capital cost for its new criteria is $35 million with an annual operation and maintenance cost of $538,000. The estimate excludes dischargers to the Atlantic Ocean, Gulf of Mexico, bayous, lakes and also assumes that essentially all plants are meeting LOT.

Utilities in Florida point to significantly higher estimates, predicting capital costs of between $24 billion and $50 billion and an estimated $0.4 to $1.3 billion dollars per year more in increased operating costs. EPA's assumption that its proposal would only impose a small, incremental increase over what Florida had proposed in 2009 explains much of the discrepancy between EPA's cost estimate and the utilities' estimates. EPA also notes that their estimate is only for inland waters, but that the utilities are including costs for all utilities statewide. Other assumptions that EPA made, noted below, also impact its costs estimate:

- The EPA analysis used as its basis a report titled "Biological Nutrient Removal Processes and Costs." Observations of the approach taken by EPA are as follows:
    - The study reviewed a number of biological nutrient removal (BNR) facilities of varying sizes and locations, principally in Maryland and Connecticut.
    - The study did not differentiate between retrofits of existing treatment plants and new facilities.
    - The study did not consider the variation in effluent performance of the facilities. Thus, information on plants that were designed to produce an effluent TN of 7 to 10 mg/l was combined with information from plants that could produce effluent TN of 3 to 5 mg/l. The report did not produce an estimated cost of a facility that was considered to meet the Best Available Technology (3 mg/l of TN, since TN is the most challenging nutrient to remove).

    - Some in the utility community have been highly critical of this report, noting that based on real-world information, it significantly underestimates costs.

- The January 2010 EPA analysis assumed that the limit of technology (LOT) for biological nutrient removal is 3 mg/l TN and 0.1 mg/l for TP. The report notes that most of the proposed in-stream nutrient standards for Florida are below the LOT, and thus it may be infeasible to meet the criteria in-stream due to technology limitations, especially if the water is dominated by wastewater discharges. The report states that if that is the case, alternative approaches would then be sought, such as a variance.
- EPA's cost estimates do not include cost to stormwater agencies because the need for incremental controls is uncertain. EPA has estimated that retrofitting storm sewer systems in the Chesapeake Bay region to address nutrient pollution would cost over $7.9 billion.
- For the purposes of the EPA cost study, it was assumed that the plants would meet the limit of technology. EPA defines the limit of technology as a BNR process and did not include microfiltration/reverse osmosis.
- The EPA cost report assumes that if a plant has an operating BNR process that is capable of treating to limit of technology, then it would not cost more money to treat to the limit of technology. This assumption ignores the reality that BNR is a biological process that will have inherent variability in its performance, and that variability will increase as the treatment goal is lowered.
- The EPA cost report does not appear to include plants that have ocean and estuary discharges. It only includes plants that discharge to lakes and streams.
- The unit costs that have been used by EPA appear to underestimate the potential costs for biological nutrient removal. For example, for a plant that is already treating for TP, the cost to add facilities to removal TN to less than 3 mg/l is given by EPA as only $0.68/gallon. The utilities in Florida estimate that the expected cost to upgrade a secondary plant to advanced treatment is about $8.20/gallon, which is more consistent with actual observed costs.

EPA's statement that most of the proposed in-stream nutrient standards for Florida are below the LOT, and thus it may be infeasible to meet the criteria in-stream due to technology limitations, especially if the water is dominated by wastewater discharges, is a significant assumption that both explains why EPA's cost estimate is so low – they assume wastewater treatment plants will not be required to treat beyond 3 mg/l TN and 0.1 mg/l for TP – but also raises the larger policy issue of whether forcing all treatment plants to LOT, when the criteria will still not be met and environmental benefit cannot be assured, is sound environmental policy. Levels below those assumed by EPA are achievable using advanced wastewater treatment with microfiltration followed by reverse osmosis. As detailed in the cost estimate developed by utilities in Florida, treatment process trains beyond advanced nutrient removal that include reverse osmosis are much more costly than EPA has estimated, however EPA has assumed that this level of protection would not be required.

NACWA believes it is counterproductive and a waste of resources to develop and implement criteria that all parties acknowledge cannot be reasonably achieved with currently available nutrient removal technology. Nothing in the CWA would allow permitting agencies to stop imposing point source controls at some defined limit of treatment technology. The current approach forces states to develop nutrient standards for TN and TP, and deal with implementation complications after the fact through variances or other regulatory options EPA discusses in the proposal. Clearly there must be a better approach that includes a common understanding of

how nutrient standards will be implemented and translated to effluent limits so that EPA and utilities can share a common baseline from which to estimate the capital and operating cost implications of the proposed rule.

EPA's Proposal Sets Goals that May Not Be Technically Achievable

EPA's proposed numeric nutrient standards for Florida raise the issue of what is really achievable in terms of meeting low levels of both TN and TP at wastewater treatment plants. EPA states in the preliminary compliance cost report for Florida (EPA, 2010) that the Agency "believes that BNR is the most reasonable and best technology for nitrogen and phosphorus removal for WWTPs." Biological nutrient removal (BNR) is typically capable of producing effluent phosphorus of approximately 1 mg/l and effluent nitrogen of 10 mg/l. Enhanced nutrient removal (ENR) with effluent filtration and larger biological reactors may produce effluent phosphorus in the range of 0.25 to 0.50 mg/l and effluent nitrogen of 4 to 6 mg/l. The most advanced nutrient removal systems operating at the maximum capability of treatment technology with multiple filtration steps or membranes, and larger biological reactors, may reduce effluent phosphorus to 0.050 to 0.100 mg/l and effluent nitrogen to 3 to 4 mg/l. EPA states that (EPA, 2010) "[f]or municipal wastewater treatment plants (WWTPs), the limit of technology (LOT) for biological nutrient removal is 3 mg/L for TN and 0.1 mg/L for TP."

EPA has not provided implementation guidance in the proposed Florida nutrient standards and so there is uncertainty as to how instream standards will be translated to effluent discharge permit limits. Comparing the very best nutrient removal facilities with the proposed Florida numeric nutrient standards reveals that effluent phosphorus may approach these instream targets with the most advanced treatment processes should they be applied end-of-pipe. However, even the most advanced nitrogen removal facilities would have effluent concentrations much higher than the proposed instream standards for Florida. The proposed Florida standards are likely to result in very restrictive discharge permit limits lower than treatment technologies are capable of achieving. NACWA is concerned about the uncertainty associated with what will be required of wastewater facilities, whether treatment technologies are available for successful compliance with the proposed standards and whether application of available technologies to achieve extremely low levels of nutrients makes sense given the significant energy and other resource demands of such technologies.

The EPA states in the preliminary compliance cost report for Florida (EPA, 2010), but not in the proposed standards, that:

> Most of the proposed criteria are at or below the LOT. For waters where the criteria are below the LOT, it may be infeasible to meet the criteria instream due to technology limitations, especially if the water is dominated by WWTP discharges. In those cases where it may be technologically infeasible to discharge at levels below LOT, the following alternatives would be considered: (1) the lakes criteria adjustment procedure or reallocation procedure of the DPV methodology for estuary protection, (2) site-specific alternative criteria, (3) restoration standards, (4) variances, or (5) UAAs and resulting changes to designated uses. Site-specific modeling would be needed to identify such instances where it is infeasible to meet criteria instream and determine appropriate effluent limits.

The proposed Florida standards have resulted in broad discrepancies in the understanding of treatment technology requirements and costs. Guidance is needed to ensure there is a common understanding of how the proposed Florida standards will be implemented. While EPA's cost analysis is based on recognition that "... BNR is the most reasonable and best technology for nitrogen and phosphorus removal," the proposed rule

dictates an instream numeric nutrient standard which requires interpretation to define compliance requirements (e.g., permit limits) that presumably could go beyond BNR. Implementation guidance is needed to address the uncertainty associated with the ability of treatment technologies to meet the proposed standards, the ability to achieve compliance with the resulting effluent discharge permit limits and the capital and operating costs and other ancillary impacts of the necessary treatment facilities.

Positive Aspects of the Florida Proposal

In the proposal, EPA has included several concepts that could provide additional flexibility in the implementation of nutrient criteria. NACWA welcomes many of these concepts, but believes that implementation flexibility alone cannot compensate for flaws in the criteria development methodology. NACWA believes that new approaches to developing and expressing nutrient criteria are needed in addition to these implementation aides.

- In the proposal, the instream protection value criteria for rivers and streams are based on an annual geometric mean. EPA proposes two expressions of allowable excursion frequency – a no more than one-in-three-years excursion frequency for the annual geometric mean and a not to exceed value for the long-term arithmetic average of annual geometric means equal to the criterion-magnitude concentration. NACWA generally supports this approach as it better reflects the behavior and potential impacts of nutrients in the water environment and believes that similar flexibility should be extended to the implementation of nutrient limits in CWA permits (e.g., annual averaging periods for permit limit compliance). However, NACWA notes that use of the arithmetic average of geometric means as a component of the allowable excursion frequency in the case of the proposed criteria for Florida actually increases the rate of false positives (more healthy waterbodies being deemed impaired) due to the incompleteness of EPA's database in Florida.

  With the exception of ammonia, nutrients have not been shown to have an immediate impact on aquatic life, and in fact the impacts normally seen as a result of nutrients are on a watershed scale. The algal response to nutrients tends to occur over a longer period of time, which makes longer averaging periods appropriate. Furthermore, it has also been shown that the lower the treatment goals are for nutrients, the greater the day-to-day variability in the biological processes. Studies of some of the most advanced wastewater treatment plants in the United States have shown ratios of reliable performance to median performance of 1.26 to 6 ($95^{th}/50^{th}$ percentile ranges). Use of a geometric mean, rather than an arithmetic mean, addresses the high variance in data sets, similar to the approach historically taken with bacteria standards to protect recreation uses.

- As proposed, EPA is allowing the development of site-specific alternative criteria (SSAC) where data and analyses show that the regional criteria may be inappropriate. NACWA supports the proposal to establish a federal procedure for adoption of SSAC in Florida that will obviate the need to formally withdraw the federally-promulgated criteria. The proposal also uses this SSAC process to allow Florida to preserve some of its existing program. If a TMDL has been done for a waterbody, Florida can use the target established in the TMDL as a SSAC. EPA also indicates that a future TMDL that establishes a target may be used as a SSAC. NACWA believes this is an essential element that will help to minimize some of the impacts of the proposed criteria on the existing Florida program. That said, NACWA

understands from its members in Florida that additional details will be needed on exactly how this process of converting TMDL target values into SSACs will work.

- EPA's proposal notes that the federal standards regulations allow states to conduct "categorical" use attainability analyses (UAAs) for certain waters, such as the canal network in Florida, that have similar hydrologic and morphological characteristics, which can be "characterized as a group and where the necessary level of protection may differ substantially from other lakes or flowing waters within the state."

- EPA's proposal recognizes Florida's existing authority to issue variances, stating that "EPA believes that the State has sufficient authority to use its adopted and EPA-approved variance procedures with respect to modification of their Class I or Class III uses as it pertains to any federally-promulgated nutrient criteria."

EPA's New Restoration Standard Concept Shows Promise, Needs More Work

For the first time, EPA is proposing a new implementation option for Florida that NACWA understands may be rolled out nationally at some point. EPA notes in the proposal that some states have already tried to provide additional time using variances in an effort to restore a particular waterbody, in lieu of changing the ultimate designated use via a UAA. EPA is seeking comment on an explicit regulatory mechanism that will preserve the current designated use and compensate from some of the limitations of the variance process. EPA proposes using a new restoration standard, in lieu of granting variances which "often lack specific milestones and a transparent set of expectations for the public, dischargers, and stakeholders," that would set certain expectations for achieving "incremental progress in a step-wise fashion" toward the ultimate water quality goal.

NACWA appreciates EPA's willingness to provide implementation flexibility, and believes the concept of a restoration standard is beneficial and could be a valuable tool in other states as well. The concept of a restoration standard would provide states with the ability to establish long-term plans for its waters and ensure the state's water quality standards continue to reflect the mandates of the CWA and EPA's implementing regulations. However, the restoration concept as described in the proposal presents a tremendous level of burden that may limit its viability. NACWA is concerned about implementation and the consequences for not meeting interim goals, but commits to working with EPA to develop an implementable restoration standard concept.

In essence, EPA is proposing that Florida go through a comprehensive UAA process, that in reality could comprise a suite of UAAs, one for each interim milestone. The resulting analysis would be more complex than a UAA used to permanently change a designated use, and could result in significant consequences for point sources. If interim designated uses and criteria are not met, and the state has not conducted a new use attainability analysis, then the use and criteria that were to apply at the end of the restoration period would become effective immediately, putting even those dischargers that had done what was required under the restoration standard out of compliance.

For point sources, additional certainty and/or off-ramps from the restoration process are needed. EPA's proposal indicates that technology-based requirements for point sources would be targeted in the first phase, noting that technology that has been successfully implemented in comparable situations and is feasible would be the focus. If wastewater treatment plants are required to install technology, whether biological nutrient

removal, enhanced nutrient removal, or limit of technology removal, during the first phase of the restoration standard, they must be given assurances that more will not be asked of them during the restoration period. These technologies are not easily implemented and will require years to install and years if not decades to recover the capital investment made.

Publicly owned treatment works (POTWs) working to meet their restoration standard obligations would need concrete milestones based on actions implemented (i.e., treatment technology installed) rather than whether the water quality criteria are met. POTWs will need assurances that they will not be deemed out of compliance if the restoration standard fails to meet an interim or final goal if the POTW has met its implementation obligations.

Again, NACWA believes that the restoration standard concept is promising, but more details and more discussion are needed before it can be successfully implemented on a large scale.

Issues for Future Discussion

Again, NACWA believes that a new approach or approaches are needed nationally on this issue and is convening a Summit of its membership to outline potential solutions. NACWA plans to use the report from the Summit to initiate a dialogue with EPA and the states on improving the water quality paradigm for nutrients.

Again, NACWA believes the following principles are essential for the development of valid and meaningful nutrient controls, ultimately leading to a new, national approach:

1. Water quality goals must be technically and scientifically defensible and technically achievable;
2. Water quality goals must be based on a demonstrated cause and effect relationship between stressor and response variables and should not be implemented until a biological impact has first been demonstrated;
3. A watershed approach using adaptive management is essential to meeting long-term water quality goals; and
4. Water quality goals should result in a net environmental benefit and consider cost and ancillary benefits in relation to effectiveness and sustainability.

These principles will guide NACWA's discussions moving forward, as will a consideration of the implementation realities we face as a nation.

Considerations for a New Approach to Nutrient Control

Implementing water quality criteria requires a complicated balancing act between two extremes of the spectrum of water use. One extreme is discharging untreated wastes to streams, similar to practices that occurred prior to enacting the CWA. The other extreme is complete protection and isolation of waterbodies with no use or impact by mankind. The reality is somewhere between these extremes.

Several factors must be considered to balance between the two extremes. For example, can a range of conditions be accepted to allow for time and data collection to discover an optimal compromise? Such an adaptive management approach may support activities that should improve water quality and the ability to

evaluate the true effectiveness of these activities. Adaptive management can distribute the associated implementation costs, making the activities more affordable, and provide better assurance that resources are spent wisely. An adaptive management approach allows for the investigation and determination of activities that are sustainable and the selection of activities to improve water quality that consider the technical feasibility to implement dependable measures. Driving all of this, an overarching watershed management plan is essential for success in ensuring that all sources and needs are balanced and addressed.

Several potential regulatory mechanisms may be used to address water quality. This includes existing regulatory mechanisms such as compliance schedules, site-specific criteria, variances, and UAAs. Other alternatives are treatment technology standards, restoration standards, and the Chesapeake Bay Program approach. Each of these approaches have their pros and cons, but independently fall short of meeting the needs of all stakeholders.

Water quality variances are an existing regulatory mechanism that can be used to reconcile unattainable water quality standards with what is possible in terms of treatment technology and affordability. Effluent performance requirements still need to be defined, but there are existing affordability reference points from EPA that bracket the potential investment required. There is, however, a substantial burden in making the case to justify receiving a variance and continuing to re-qualify on a periodic basis up to 20 years without assurance of a satisfactory outcome.

EPA's newly proposed restoration standard approach appears similar to a water quality variance. The restoration standard approach allows for an extension of full compliance by phasing for eventual attainment of water quality standards. The restoration standards concept appears to place greater emphasis on beneficial use attainment in the watershed in phased milestone steps, and less emphasis on the exceptions qualifying for variances, such as economic hardship. Restoration standards would have a substantial burden in building the justification that would include demonstrating that standards are not attainable, and defining interim milestones, beneficial uses, and "maximum feasible progress." EPA appears to be framing restoration standards to consider broader watershed issues including nonpoint source controls and innovative and flexible approaches.

None of the existing regulatory mechanisms, however, cover all of the factors that NACWA believes are needed for an effective national program. And these mechanisms only address implementation, not the underlying criteria development process. Above all, none of the approaches address equity or require nonpoint source reductions. NACWA will be evaluating the current approaches to nutrient goal development and the various implementation mechanisms in advance of and during its Summit this summer. NACWA has already been exploring some of the innovative approaches being considered by states and other stakeholders.

States Are Trying New Approaches to Ensure Criteria Reflect Biological Conditions

The Ohio Environmental Protection Agency is currently in the process of developing nutrient criteria for streams and inland lakes in the state. To develop the criteria, change points in data relating benthic chlorophyll *a* to nutrients are being compared to desired objectives for macroinvertebrate and fish community health and dissolved oxygen levels and diel swings. It has been indicated that a weight-of-evidence approach would be the most appropriate way of implementing these nutrient criteria. Biological impairment would be the first trigger before the nutrient criteria would apply. For sites in which the biological community is not in

attainment, nutrient concentrations, DO swings, and chlorophyll *a* concentrations would all be evaluated to determine whether nutrients are contributing to the impairment. If nutrients are determined to be a source of impairment, the developed numeric criteria would then need to be met and may be used to set effluent limits. NACWA believes that approaches like those being pursued by Ohio EPA may provide a more scientifically defensible and reasonable approach to implementing nutrient water quality programs.

Nutrient Assimilation Zone Concept

Implementation flexibility for nutrient controls is essential and some stakeholders are working on a nutrient assimilation zone (NAZ) concept to represent an area or volume in a waterbody where both traditional dilution and nutrient assimilation is allowed. Nutrient goals could be exceeded within the proposed NAZ but would be met at its boundaries. Dilution within mixing zones is presently allowed by regulation to comply with acute and chronic water quality criteria for toxics. Unlike toxics, nutrient effects are subtle and tend to occur along continuums as opposed to "bright line" thresholds. As such, effective management of nutrient discharges should allow for nutrient assimilation in addition to traditional dilution.

Currently available numerical models have the potential to estimate expected responses in various end points (*e.g.*, algal growth) associated with various levels of treatment. The results of such models could be applied to determine appropriate site-specific controls. This proposed modeling approach is consistent with the findings of the SAB review that advocated for greater use of mechanistic tools rather than a reliance on simple statistical approaches. NACWA plans to further explore the NAZ concept as well as other approaches designed to improve the implementation of nutrient controls during its Summit this summer.

Conclusion and Recommendations for Next Steps

NACWA believes the situation in Florida is symptomatic of a larger underlying problem with EPA's current approach to addressing nutrient over-enrichment, and underscores the need for new approaches. Again, NACWA offers the following recommendations for moving forward with a more reasonable and scientifically valid program for addressing nutrient enrichment:

1. If EPA continues to pursue the federal criteria for Florida as proposed, the Agency should draft and submit for public comment implementation guidance before the criteria are promulgated;
2. Nationally, EPA must change its existing framework that necessitates the establishment of numeric criteria for TN and TP, which may not be appropriate for all waters;
3. EPA should better recognize and approve new and innovative approaches for expressing nutrient water quality goals currently being considered by some states;
4. Future water quality criteria development work must be supported by information on biological response to ensure criteria will actually protect designated uses;
5. Additional flexibility, both in criteria development and in permitting (e.g., averaging periods), to account for the unique behavior of nutrients must be provided;
6. Ultimately, EPA, the states and remaining stakeholders must work to establish a new, more holistic approach to addressing water quality problems that includes meaningful and equitable controls and implementation requirements for all nutrient sources.

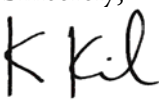
NACWA believes a new dialogue must begin to outline more reasonable approaches to nutrient control and that additional scientific and technical work is needed, including in the following areas, to arrive at supportable outcomes:

- Continue research to better define biological impacts of nutrients in different waterbodies;
- Improve scientific understanding of the fraction of TP and TN that is not bioavailable;
- Conduct a comprehensive evaluation of the social, economic and environmental impacts of different concentrations of nutrients in different types of waterbodies versus the social, economic and environmental impacts of the different levels of nutrient treatment.

Again, NACWA appreciates the opportunity to comment on the proposal which will have impacts on NACWA's members in Florida and implications more broadly for the nation. NACWA will work with EPA as it convenes its Summit and looks forward to a collaborative dialogue with EPA, the states and other stakeholders on the issue.

Please contact Chris Hornback at 202-833-9106 or *chornback@nacwa.org* if you have any questions or would like to discuss NACWA's comments further.

Sincerely,



Ken Kirk
Executive Director

ENCLOSURES

REFERENCES

HydroQual, 2000. Review of Nutrient Criteria Technical Guidance Manual: Rivers and Streams.

Science Advisory Board, Ecological Processes and Effects Committee, 2010. *Advisory Report – For Quality Review*. January 8, 2010.

U.S. Environmental Protection Agency, 1995. *Water Quality Guidance for the Great Lakes System -- Supplementary Information Document.*

U.S. Environmental Protection Agency, 2000. *Ambient Water Quality Criteria Recommendations - Information Supporting the Development of State and Tribal Nutrient Criteria Lakes and Reservoirs in Nutrient Ecoregion XII.*

U.S. Environmental Protection Agency, 2001. Memorandum from Geoffrey Grubbs, Director, Office of Science and Technology to Water Directors, Regions I-X, et. al. *Development and Adoption of Nutrient Criteria into Water Quality Standards*. November 14, 2001.

U.S. Environmental Protection Agency, 2009. Benjamin Grumbles, Assistant Administrator, EPA Office of Water, to Michael Sole, Secretary, FDEP, January 14, 2009.

U.S. Environmental Protection Agency, 2010. *Preliminary Estimate of Potential Compliance Costs and Benefits Associated with EPA's Proposed Numeric Nutrient Criteria for Florida*. January 2010.

Appendix 1

Review of Nutrient Criteria Technical Guidance Manual: Rivers and Streams

Review of

Nutrient Criteria Technical Guidance Manual: Rivers and Streams

USEPA-822-B-00-002 (July 2000)

The Technical Guidance Manual is a good resource for defining certain elements with which States/Tribes can develop nutrient and algal criteria for rivers and streams. The document defines many approaches for first classifying the physical environment and stream types, then choosing variables for a water quality assessment, then data analysis for criteria setting followed by implementation. Depending on resources, data gathering prior to criteria setting and monitoring and, likewise, re-evaluation afterward may occur. While the flexibility incorporated in this document is commendable, the incorporation of less-rigorous approaches may lead to criteria setting by some parties that is not scientifically defensible, a stated goal of the document. The comments below address suggested approaches within the guidance document which fall short of providing credible means of setting criteria.

The response of a water body to nutrient inputs is highly variable. Many factors—hydrology, geology, stream morphology, degree of canopy cover, land use, grazing organisms — influence a particular water body's assimilative capacity for dissolved nutrients. Thus, trophic status is not uniquely determined by nutrient load. As a result, the use of a single numeric nutrient criterion applied to a given region is likely to be under- or overprotective for a given water body. Scientifically defensible regulation of nutrients should focus on prevention or control of the undesirable effects of nutrient over-enrichment rather than on absolute nutrient limits.

In general, the establishment of criteria for so-called "primary" or "causal" variables such as nutrient element concentrations represents a departure from the modern process of developing standards based on definitive cause-and-effect relationships. For example, there is no criterion for biological oxygen demand (BOD); rather, there are waterbody-type-specific criteria for dissolved oxygen (DO). This is because the level of BOD itself does not determine the level of impairment within an individual waterbody, although indeed it may contribute to impairments. The proximate factor leading to the loss of biological integrity in an ecosystem is a depletion of DO, with chronic and acute effects on specific kinds of desirable or keystone organisms. Levels of BOD can be used to develop permit limits in a system-wide waste load allocation so as to protect the DO criterion, based on site-specific evaluations of the interaction of BOD with other processes which determine DO concentrations. It is this philosophical departure from effects-based criterion setting which is most disagreeable and which creates a fundamental problem of scientific credibility for the regional nutrient

criteria setting process. It is strongly suggested that the process be revised to specifically address cause-and-effects factors.

The present guidance document contains many approaches for what may be called exploratory data analysis. These procedures are designed to uncover patterns in the primary variables from regionally aggregated data. However, the inability of such analyses to establish causal effects places constraints on their efficacy. The suggested procedures can be used to locate water bodies which may be susceptible to impairment. Statistical analyses based on regional or national data should be used only as screening tools to initiate further site-specific investigations, not for establishing criteria. The distinction between criteria versus screening-level warning flags is legally significant and needs to firmly established in the technical guidance documents.

In order to demonstrate the disconnect between primary variables (i.e., nutrients) and specific ecosystem (or use) impairments, several areas where we feel the Technical Guidance Document fails to establish scientifically credible interpretations will be discussed. Most of these issues could be resolved if it is recognized that the suggested solutions for gathering and evaluating data are extremely effective means for establishing screening values for flagging potential impacted systems rather than for establishing numeric nutrient criteria. These comments address the following components or approaches contained in the Technical Guidance Document:

(1) the frequency distribution approach to criteria setting,
(2) the ecoregional approach to classification,
(3) interpretations of relationships between dissolved nutrients or total nutrient element pools, e.g., total nitrogen (TN) or total phosphorus (TP), and algal biomass based on nationwide data, and
(4) a model for criteria development.

Frequency Distributions and Criteria Setting

In Section 7.5 "Methods for Establishing Nutrient and Algal Criteria," three approaches are outlined for defining a "reference reach" for establishing nutrient concentrations or algal biomass levels which would protect a stream's "natural biological integrity." While one relies on establishing reference conditions based on "best professional judgement (BPJ), the other two approaches use frequency distributions of available data for the primary variables (nutrients, algal biomass, water clarity as secchi depth, turbidity, or a measure of total suspended solids). One of these two methods suggests criteria-setting at the 75th percentile of a data population from reference streams (presumably

identified via BPJ), while the other sets a criterion level at the 5th or 25th percentile of data from all streams within an ecoregion. Lastly, it is suggested that a criterion value might be set at some level in between these reference and all-stream percentiles.

There is a strong caveat to the manager about using this approach at the end of this section. The document acknowledges the ambiguity and difficulty in picking a criterion level from this approach ("A single criterion forces the manager to make decisions about the number of streams that will be in unacceptable condition....", p. 97). In fact, such decision-making will have no basis in whether a given stream in fact has had its biological integrity compromised by a criterion level established from some lumped distribution of nutrient data from many streams and locations. Uncertainties about the relative distribution of data from reference versus impacted sites within the database may lead to criteria which are either too stringent or not stringent enough. More importantly, evaluation based on frequency distributions of primary variable data from a variety of streams even within some defined ecoregion simply does not contain information about the causal relationships between those variables and the potential for impairment of biological integrity in any given stream or reach. Even in light of the document's stated caveats, managers may view this is the simplest, least-costly approach to criterion setting, without a clear understanding of the processes leading to a given stream's susceptibility for impairment or ability to assimilate nutrients without impairment.

Ecoregional Classification of Waterbodies

The classification of water bodies in a hierarchical manner is presented so that managers have a means to compare data and make extrapolations among streams with common features and presumed similar functionality. Approximately half a dozen geologically, hydrologically, and/or biologically based schemes are presented. While Omernik's ecoregional classification scheme represents a starting point in this process, the final procedures for establishing meaningful sub-regions is left to the States and Tribes. Examples provided suggest several important interactions of factors which influence the primary nutrient, biomass, and light variables of concern, including underlying geology, land use, substrate type, and dominant primary producer (e.g., Table 1, p. 21). In the end, the determinants of a given stream's trophic state (Table 2, p. 27) may be either natural or human-induced, with little means of determining the causality. Because such classification schemes impact choice of monitoring sites, reference sites, and criterion-setting, this lack of a causal link between a stream's primary variables and the underlying processes which determine a stream's trophic state, incorrect inferences may be drawn about what nutrient concentration or algal biomass level is protective of a particular stream's designated and beneficial uses.

Interpretations of Relationships Between Nutrients and Algal Biomass

The document states that predictive relationships between nutrients and algal biomass are required to identify critical or threshold nutrient levels that produce nuisance algal blooms (p. 76). However, the simplest approach of regressing algal biomass versus nutrient element concentration (as either TN or TP), as stated, provides only moderate explanatory power, accounting for one-third or less of the observed variance in periphyton biomass. Suggested regression equations for suspended algae yielded somewhat better correlations, up to 73% of the variability, if the catchment area for the stations is known. As the document goes on to state, "critical and highly variable factors other than nutrient—shading, [substrate], scour, water level fluctuations..., grazing intensity—have major effects on algal biomass" (p. 76-77). In fact, as shown in Figure 7 of the document (p. 78), at 20 ug/L TP, mean algal biomass as chlorophyll a varied by two orders of magnitude while at 500 ug/L TN, mean biomass varied by a factor of 25. Those nutrient values are within suggested criteria ranges for nutrients presented in Table 4 (p. 101). In an investigation of 26 sites in seven streams, much of the control on periphyton biomass was influenced by macroinvertebrate grazers, riparian shading, and substrate (Welch et al., 1992). Without consideration of the multiple factors that influence algal biomass and species composition and ultimately influence dissolved oxygen and pH levels, nutrient-based criteria cannot be established with scientific credibility for a given stream environment.

The "Tennessee Ecoregional Nutrient Criteria" case study presented in Appendix A does serve as a good example of the use of approaches recommended in the guidance document to data analysis and the process of establishing numeric nutrient criteria. Unfortunately, it also shows the difficulty of actually ascertaining impairment or degrees of protection offered by even sub-ecoregionally applied numeric criteria. According to the Tennessee plan, "[s]treams with nutrient levels higher than the reference stream database range [at the 90th or 75th percentile, to be determined] will be considered in violation of the narrative [nutrient] criteria" (p. A-7). Such streams will be placed on the state 303(d) list. Yet, in an analysis of data relationships, nutrients were extremely poor predictors of response variables, such as total organic carbon (TOC) or turbidity. Thus, nutrients alone were not indicators of impaired uses.

Further analysis comparing data from the national database with recent monitoring data from identified reference streams demonstrated a potential for bias in the national database. EPA draft Nutrient Aggregate Ecoregion IX data, covering western Tennessee, had TP values ten-fold that found at the reference sites. The authors postulated that the national database stations may either have been biased by locations with high-phosphorus content soils or with sampling data targeted to quantity worst-case nutrient loading events. Without sufficient meta-data or local knowledge to stratify the

data sets for such biases, procedures such as analysis of frequency distributions are likely to generate criterion values which may be under- or over-protective.

If, instead of using these empirical tools as means of establishing criteria, for which they have little power, they were used to establish screening values for particular ecoregional stream types, the guidance would be more acceptable.

Metals Criteria: A Model for Criteria Development.

Over the past two decades, national criteria for certain metal toxicants has progressively moved from a single numeric value to varying stages of site-specific values based on local environmental factors, such as water hardness, pH and dissolved organic carbon concentration, that influence a metal's bioavailability (e.g., Renner, 1997). Specific laboratory and statistical procedures have been developed to quantify effects in view of environmental factors. Lessons learned from metals criteria development should be applied to nutrient criteria development.

The importance of site-specific factors in determining appropriate pollutant limits is not unique to nutrients. The impacts of some toxics are also modified by receiving water characteristics. For example, the toxicity of metals, including Cd, Cu, Ni, Pb and Zn, to aquatic organisms typically range over several orders of magnitude for a single organism (Chapter 5; Meyer, 1999). In the natural environment, various processes that modify metal toxicity have been identified. In response to these problems the US EPA has allowed the use of site-specific adjustments to the ambient criteria for metals (e.g., the water-effect ratio procedure - WER; United States Environmental Protection Agency, 1994). The use of WER procedures to modifiy numeric criteria for metals, however, requires time-consuming and expensive bioassay testing. Recently, an alternative approach has been developed using a model of metal bioavailability and toxicity (Di Toro et al., 2000; Santore et al., 2000). The Biotic Ligand Model (BLM) uses a mechanistic understanding of the effects metals have on aquatic organisms and the way these effects are modified by the physical and chemical characteristics of a receiving water body. The use of a mechanistic model for determining criteria provides a scientifically defensible approach for considering the complex site-specific factors that determine the impact of metals in the environment.

The approach for establishing nutrient criteria in receiving water bodies should address many of these same concerns. A modeling approach for setting nutrient criteria can be designed to incorporate complex site-specific factors. Furthermore, a mechanistic framework will provide a direct linkage of nutrient loads with ecosystem response. This linkage will allow a proposed criteria to be based on the prevention or control of harmful effects such as low dissolved oxygen concentrations or trophic status

rather than on nutrient concentrations per se. This type of mechanistic framework can explicitly consider the natural variability in the response of receiving water bodies to nutrient inputs and, therefore, is more scientifically sound than a regional numerical value determined from percentile values.

References Cited

Di Toro, D.M., H. Allen, H. Bergman, J. Meyer, P. Paquin, R. Santore. 2000. "Biotic ligand model of the acute toxicity of metals. I. Technical Basis." *Environmental Toxicology and Chemistry*. accepted for publication.

Meyer, J. 1999. *Literature review on the bioavailability and toxicity of selected metals in the aquatic environment relative to natural water chemistry*. International Lead-Zinc Research Organization, Research Triangle Park, NC.

Renner, R. 1997. "Rethinking water quality standards for metals toxicity." *Environmental Science and Technology* 31:466A-468A.

Santore, R.C., D.M. Di Toro, P.R. Paquin. 2000. "A biotic ligand model of the acute toxicity of metals. II. Application to acute copper toxicity in freshwater fish and daphnia." *Environmental Toxicology and Chemistry*. accepted for publication.

United States Environmental Protection Agency. 1994. *Water Quality Standards Handbook.* United States Environmental Protection Agency. Office of Water, Washington, D.C.

Welch, E.B., J.M. Quinn, C.W. Hickey. 1992. "Periphyton biomass related to point-source nutrient enrichment in seven New Zealand streams." *Water Resources* 26:669-675.

Appendix 2

Info Tech Critique
of
EPA's Proposed Numeric Nutrient Criteria for Streams

Introduction and Executive Summary

Info Tech has performed a statistical review of the nutrient criteria for total nitrogen and total phosphorus in streams as provided in Chapter 2 of the Technical Support Document for U.S. EPA's proposed Numeric Nutrient Criteria for FL Inland Surface Fresh Waters. Specifically, we have addressed the issues of (1) whether the data is sufficient to reliably generate an in-stream protective value, (2) whether all sources of variability have been taken into account in determining the criteria, and (3) whether the false positive rates resulting from adoption of the EPA's proposed criteria comport with widely accepted statistical and scientific principles.

As detailed herein, we find that EPA's proposed criteria fail on all three counts. The data employed by EPA is insufficient to provide reliable estimates of threshold criteria for water quality. Data availability at most sites utilized in EPA's derivation of proposed standards are particularly under-representative temporally, with 85% of sites having no more than two annual observations. EPA's "solution" is to aggregate site data over all time (i.e. averaging over months and years for a given site), which disguises rather than solves the lack of temporal coverage in the data.

As a result of the insufficient data and EPA's over-aggregation, its proposed criteria fail to reflect proper estimates of variability inherent in the water quality data. Indeed, the insufficiency of the data makes it virtually impossible to obtain reliable statistical estimates of these important components. This means that factors which contribute to the change in nutrient levels have not been properly captured and considered in determining a threshold value.

The net result of these data and methodology inadequacies is that EPA's criteria are arbitrary, and will result in declaring a statistically and scientifically unacceptably high percentage of compliant sites out of compliance ("false positives").

1. Data are insufficient to generate reliable criteria.

The instream protective criteria for streams proposed by the EPA rely upon insufficient data necessary to obtain a reliable estimate of a protective threshold. Fully 70% of the sites utilized by the EPA have data for only one year and another 15% have data for only two years during the 19 year period covered by the data. Thus, 85% of the sites cover less than 11% of the sampling period. As a result of having such limited time coverage, the temporal (year-to-year) variability cannot be reliably measured for the great majority of the sites, and therefore this critical component is missing from

EPA's threshold calculations. That is, temporal effects in the large majority of values are confounded with spatial effects. Instead, the EPA simply ignores the year-to-year variability by utilizing sites aggregated over years, which, for 70% of the sites, is a single observation in a single year; and for another 15% of the sites, the EPA average consists of two observations. This does not produce statistically defensible criteria that can be used to judge whether an annual site mean is or is not out of compliance.

The EPA also proposed an alternative methodology developed by the FDEP with a similar insufficient data issue. It utilizes data aggregated at the water body-year level by averaging over sites and sampling events. With sufficient data, a statistical method that considers the spatial (site-to-site) and temporal (year-to-year and month-to-month) variability could be used to develop a reliable criterion. However, using aggregated data that ignores the variability inherent in the data will not reliably measure nutrient levels needed to regulate water quality at a given site. The EPA claims "to reduce the effect of low sample sizes, while capturing the temporal variability and maintaining the maximum amount of data being considered, a minimum of four samples per year in each WBID was established as the appropriate requirement for data to be included in further analysis of the regional distribution". (*Page 2-38, Chapter 2, EPA Technical Support Document.*) Aggregating data across months does not "capture the temporal variability," since the change in nutrient levels from month-to-month, e.g. a seasonal effect, cannot be estimated. Since the variation in nutrient data is due to spatial, temporal, and site-specific effects, model-based estimates of these sources of variation are needed in order to make valid assessment of variation needed for compliance criteria.

A critical assumption of any proposed criterion is that the sample of water bodies on which it is based is truly representative of the water bodies to which the criterion is to be applied. The inadequate sample utilized by the EPA cannot reasonably be assumed to represent even the sampled water bodies, much less be meaningfully applied to all or even a significant portion of other streams on a regional or statewide basis. Since the vast majority of the sites were sampled for only one or two years, the data are not sufficient to enable a reliable estimate of the variability. Statistically defensible criteria must reliably measure the extent to which each of the relevant components affect the change in nutrient levels. The sample employed by the EPA is neither representative of the population of water bodies, nor is it sufficient to estimate reliably the contribution of relevant factors to the levels of the water quality parameters.

2. <u>"False Positive" rates are unacceptably high for the proposed criteria</u>

Regardless of which set of standards are adopted – EPA or FDEP -- the chance of declaring a site out of compliance when it is truly in compliance is a critical attribute of any test criteria. The probability of incorrectly declaring a site out of compliance is referred to as the "false positive rate" for the compliance criteria. False positive rates are routinely calculated and published for medical screening tests, where a "false positive" occurs when the test incorrectly indicates an individual has a disease and

treatment is therefore unnecessarily applied. The goal is to develop tests that have low false positive rates, particularly when the treatment is invasive and expensive. Similarly, the false positive rates for water quality compliance criteria are important in evaluating the likelihood of undertaking extensive, expensive, but ultimately unnecessary, remedial actions.

The EPA Instream criteria are based upon the 75th percentile of an inadequate database. Because the data are insufficient to generate reliable criteria, precise quantification of the true false positive rates is impossible. Nonetheless, an approximation can be calculated to illustrate the fact that use of the 75th percentile leads to unacceptably high false positive rates.

As shown in the following table, even using the EPA's estimated means and standard deviations, and their proposed criteria based on the 75th percentile, false positive rates range from 20% to 36% across the four regions and two water quality parameters.

EPA Instream Protective Criteria

False Positive Rates Assuming Distributional Statistics in Table 2-4 of TSD Chapter 2

Parameter	Region	Mean LN	Std Dev LN	75th Percentile	False Positive Rate
TN	Bone Valley	0.309	0.412	1.798	20%
(mg/L)	Panhandle	-0.724	0.788	0.824	27%
	Peninsula	-0.24	0.633	1.205	24%
	North Central	-0.096	0.722	1.479	26%
TP	Bone Valley	-0.752	0.667	0.739	25%
(mg/L)	Panhandle	-3.881	1.081	0.043	33%
	Peninsula	-2.879	0.957	0.107	31%
	North Central	-1.847	1.22	0.359	36%

False Positive Rates Include the Proposed Frequency and Duration Tests:

No more than one in three years excursion frequency for the annual geometric mean.

Long-term arithmetic average of annual geometric means to not exceed the 75th percentile.

The long-term average is assumed to be a seven year period.

Rates are generated based on a simulation of 10,000 site-years.

The implication of a 33% false positive rate is that one in three sites that are actually *in* compliance will be falsely declared *out* of compliance. This unacceptably high rate of error is further exacerbated when many sites are tested. For example, if five sites that are *in* compliance are tested, the probability that at least one will be declared *out* of compliance jumps to 86%, and if ten in-compliance sites are tested, the chance that at

least one will be falsely declared out of compliance is almost certain, at 98%. As unacceptable as these rates are, the fact is they may be even higher, since all these calculations are based on the assumption that the criteria were calculated properly using sufficient data, which is a demonstrably incorrect assumption.

The false positive rates for the DEP criteria are significantly lower for most regions, since for all but one region-water quality parameter they are based on the 90th percentile. As shown in the following table, the rates are generally less than 5%, with the Bone Valley TP rate higher at 22% due to the DEP's use of the 75th percentile for that region.

EPA Alternative In-stream Protective Criteria Developed by the FDEP

False Positive Rates Assuming Distributional Statistics in Table 2-16 of TSD Chapter 2

Parameter	Region	Mean LN	Std Dev LN	Recommended Percentile	False Positive Rate
TN	Panhandle	-0.760	0.450	0.82	3%
(mg/L)	NE-NC-Pen-BV	-0.030	0.450	1.73	3%
TP	Panhandle	-3.662	0.768	0.069	4%
(mg/L)	Northeast	-3.205	0.714	0.101	4%
	North Central	-1.973	0.656	0.322	3%
	Peninsula	-2.746	0.464	0.116	3%
	Bone Valley	-1.211	0.491	0.415	22%

False Positive Rates Include the Proposed Frequency and Duration Tests:

No more than one in three years excursion frequency for the annual geometric mean.

Long-term arithmetic average of annual geometric means not to exceed the 75th percentile.

The long-term mean is assumed to be a seven year period.

Rates are generated based on a simulation of 10,000 site-years.

A false positive rate of 5% or less is a significant improvement over those using the EPA's proposed criteria. However, this rate is applicable only if tested against water body annual means and assuming the FDEP have accurately estimated the water-body level statistical distribution of the nutrient parameter. The false positive rates will increase if the limits are applied using properly estimated site level statistical characteristics.

Regardless of which set of standards are adopted – EPA or FDEP – the chance of declaring a site or water body out of compliance when it is truly in compliance is a critical attribute of any test criteria. The criteria must be applied at the same level at which it is calculated, which for the FDEP data is to water body annual means, and not site level annual means, as we understand is proposed by the EPA.

3. EPA's use of the 75th percentile is not a scientifically defensible approach to solve an inadequate data problem.

The false positive rates associated with EPA's use of the 75th percentile based on inadequate data are not defensible. From a statistical perspective, the process of determining whether a particular site is in compliance is a "test of hypothesis." In scientific applications of statistical methods, "tests of hypotheses" are used to evaluate which of two competing hypotheses are supported by the data. As it applies to EPA's proposed water quality tests at a particular site, the two hypotheses are "Site is compliant" vs. "Site is not compliant." Statisticians recognize that when data are utilized to test hypotheses, two types of error can occur: in this case a Type I Error occurs if the site is declared not compliant, when in fact it is compliant. We have been referring to the Type I Error as a false positive. A Type II Error occurs if the site is declared complaint, when in fact it is not compliant. This is often referred to as a "false negative."

A generally accepted scientific principle among statisticians and scientists applying statistical methods is that the test should be conducted in such a manner that ensures the Type I Error rate will be no greater than10%, and tests of hypotheses are predominantly conducted using no more than a 5% false positive rate. When the EPA establishes the 75th percentile as the "not to exceed" level, it is accepting Type I Error rates of 25% (or greater), much higher than generally accepted scientific levels.

The 75th percentiles presented in Table 2-4 are estimates of the true 75th percentiles and do not account for sampling variation in the mean and standard deviation estimates. Doing so, using tolerance limits or confidence intervals for the percentiles, would result in a multiplier greater than 1. For example, the multiplier for the Bone Valley estimated percentiles is about 1.25, giving an upper limit on the 75th percentile for TP of 0.924 instead of 0.739. Thus, the true false positive rate would be even larger than the values shown in the table in the previous section.

The ultimate goal in statistical hypothesis testing is to minimize the rates of both Type I and Type II Errors. This goal is accomplished by first establishing the Type I Error rate at a scientifically acceptable level -- typically at 5% -- and then selecting a sample of sufficient size to ensure that the Type II Error rate is also small. The EPA's use of the 75th percentile in conjunction with inadequate data will result in unacceptably high levels of both the "false positive" and "false negative" rates. The appropriate approach is first to collect sufficient data upon which to base a compliance test, and then to determine the criteria for the test based on sound scientific principles. The current proposal satisfies neither of these conditions.

4. Use of the Arithmetic Average of Long-Term Geometric Means Increases the False Positive Rates.

The false positive rates are needlessly increased since the duration test is based on an arithmetic average of geometric means. Due to the skewed nature of the water quality data, a straight average of geometric means will skew the result high and can

increase the false positive rates. For the proposed in-stream criteria as given in Table 2-4 of Chapter 2, the EPA Technical Support Document, the false positive rates are as much as 20% higher when the long-term average is based on the inappropriate arithmetic rather than the geometric mean. Of course, as discussed above, an exact quantification of this rate is not possible due to data insufficiencies. It is known however, that the more skewed the distribution, the greater the increase in the false positive rate due to the inappropriate use of the arithmetic mean.

5. All but one of the Alternative Duration and Frequency Expressions Do Not Sufficiently Reduce the False Positive Rates

The EPA requested comment on the following alternative duration and frequency expressions:

1) a criterion-duration expressed as a monthly average or geometric mean,
2) a criterion-frequency expressed as meeting allowable magnitude and duration every year,
3) a criterion-frequency expressed as meeting allowable magnitude and duration in more than half the years of a given assessment period,
4) a criterion-frequency expressed as meeting allowable magnitude and duration as a long-term average only, and
5) an expression of the criteria in terms of an arithmetic average of annual geometric mean values based on rolling three-year periods of time.

Of all the above alternatives, only the third one, a criterion-frequency expressed as meeting allowable magnitude and duration in more than half the years of a given assessment period produces statistically reasonable false positive rates using the EPA criteria. If a seven year period is assumed and no more than three years are allowed to be in exceedance, then the false positive rate using this method is 7%. This calculation is based on the EPA's distributional statistics and limits in Table 2-4, Chapter 2, EPA Technical Support Document.

The other alternatives produce false positive rates that are above statistically acceptable standards. We assume that the first alternative expression of a monthly geometric mean would translate to monthly testing of a site. If a site is tested monthly, the false positive rate for the year skyrockets. For example, if the false positive rate for a site tested in a given month is 10%, then the chance of declaring that site out of compliance when it is truly in for at least one of the 12 months in a year is 72%. The second alternative -- a site meeting the 75th percentile every year -- would by definition produce a 25% false positive rate, well above widely accepted statistical standards. The fourth alternative -- based on a long-term average only -- produces false positive rates from 10%-31%, assuming a seven year average. The fifth alternative -- a rolling three-year period -- produces false positive rates from 18% to 29%, depending upon the parameter and region. In addition, this criterion is based on an inappropriate arithmetic average of annual geometric means, the flaws of which are detailed in Section 4. The false positive rates of each of these alternatives were estimated by a simulation covering 10,000 years, utilizing EPA's distributional statistics as shown in Table 2-4 of

the Technical Support Document. If a criterion were to be proposed that combines two or more of these alternatives, the resulting false positive rate will be higher than that of any of the component alternatives applied separately.

In summary, all alternatives first require sufficient data to obtain a statistically reliable estimate of the water quality threshold for each parameter. Assuming the threshold is reliably estimated – an unwarranted assumption for the EPA's stream analyses – only the third alternative, which calls for a compliance test based on exceedances in more than half of the years in an acceptance period, has the potential for statistically acceptable false positive rates. It does not necessarily follow, however, that improving the false positive rate would render the criteria biologically relevant or appropriate for their intended purpose in the proposed rule.

6. Criteria for Response Variables

Our criticism of the proper methodology to develop a statistical based limit applies to stressor variables as well as response variables, such as chlorophyll a. Model-based estimates of the sources of variation are needed to construct a reliable protective threshold.

7. Conclusion

The in-stream protective criteria for streams proposed by the EPA rely upon insufficient data necessary to obtain a reliable estimate of a protective threshold. The EPA ignored the year-to-year variability by utilizing stations aggregated over years, which, for 70% of the stations, is for only one year. Even assuming the criteria are calculated reliably -- which they are not -- the false positive rates exceed statistically and scientifically unacceptable levels, which, if the criteria were adopted, would lead to large numbers of incorrect and costly "out of compliance" decisions. The EPA provided an alternative methodology developed by the FDEP, which is also based upon insufficient data. It too uses aggregated data, but collapsed to the water body level. Although the FDEP use of the 90^{th} percentile will in theory result in more reasonable false positive rates, the insufficient data still renders the criteria unreliable. In addition, even with sufficient data the criteria must be applied at the same level at which it is calculated, which for the FDEP data is to water body annual means rather than site level annual means.

The net result of these data and methodology inadequacies is that EPA's criteria are arbitrary, and will result in declaring a statistically and scientifically unacceptably high percentage of compliant sites out of compliance ("false positives").