

RESPONSE TO COMMENTS FOR
LISTING METHODOLOGIES FOR IDENTIFYING IMPAIRED SURFACE
WATERS FOR LISTING ON MARYLAND’S 303 (D) LIST

ADMINISTRATIVE PROCEDURES ACT

Introduction

The Maryland Department of the Environment solicited comments on the Listing Methodology for Listing Waters on Maryland’s 303(d) List from May 21, to July 31, 2001. A list of commentators, their affiliation, and the date they submitted comments.

General

List of Commenters

Author	Affiliation	Date
Cy Jones	Washington Suburban Sanitary Commission	6/18/01

Comments and Responses

Comment 1. Gov’t Code Ann. § 10-101(g). Rulemaking cannot be accomplished by the device of soliciting comment on draft decision criteria.

MDE Response: The Department respectfully disagrees with the commentator’s characterization of these methodologies as regulations. MDE has adopted by regulation both numerical and narrative water quality criteria. The methodologies at issue will guide MDE in determining whether these criteria are being met in Maryland’s waters, but they are not intended to change or serve as independent water quality criteria. Accordingly, the methodologies are not regulations within the meaning of the Maryland Administrative Procedures Act (APA).

Importantly, moreover, the methodologies in and of themselves will not have a direct impact on the public. Rather, they will be used by the Department to collect and analyze water quality data. These technical analyses may result in the determination that a waterbody is impaired and, subsequently, that water quality-based effluent limits are necessary for a given discharge. When a permit is issued for such a discharge, the water quality-based effluent limits and other permit terms would, as always, be subject to contested case hearing and judicial review provisions of

the APA. By inviting public participation on its technical methodologies in advance, the Department has sought to expand upon, not limit, the public's opportunity to participate in the listing and, indirectly, permit development processes.

Toxics

List of Commentors

Author	Affiliation	Date
Patricia Gleason and Elaine Harbold	U.S. EPA-Region III	6/15/01
Cy Jones	Washington Suburban Sanitary Commission	6/18/01
Steve Stewart	Baltimore County DEPRM	6/21/01
Jim Stine	Constellation Energy Group	7/31/01
Jules J. Loos	Mirant Mid-Atlantic	7/31/01

Comments and Responses

Comment 1. According to the methodology, the Maryland Department of the Environment (“MDE”) will consider all existing and readily available data, including independent studies by external sources. *Proposed Methodology for Determining Impaired Waters by Chemical Contaminants for the Maryland § 303(d) List* (March 20, 2001) (“Methodology”), p. 5. Not all readily available data are necessarily relevant or appropriate for making listing decisions. In some cases, an exceedance of a water quality standard (“WQS”) is due to natural background conditions (such as high aluminum content in non-contaminated sediments, or low dissolved oxygen in portions of wetlands or deep estuarine waters) or physical alterations not related to pollutant discharges (such as higher temperatures in dammed waterways). MDE’s listing methodology should specifically state that such exceedances should be excluded from data being used for listing decisions. Likewise, exceedances within allocated mixing zones should be excluded, since those exceedances are authorized by the state and already determined not to interfere with designated uses. Florida has adopted these exclusions in its new listing regulations, which state:

“It is not the intent of this chapter to include waters that do not meet water quality criteria solely due to natural conditions or physical alterations of the waterbody not related to pollutants. Similarly, it is not the intent of this chapter to include waters where designated uses are being met and where water quality criteria exceedances are limited to those parameters for which permitted mixing zones or other moderating provisions (such as site-specific alternative criteria) are in effect”. – (Fla. Admin. Code 62-303.100(2). Maryland’s methodology should incorporate these exclusions).

MDE Response: The Department "considers" and evaluates all data. If an apparent exceedance of water quality standards occurs that can justifiably be attributed to natural conditions, the Department will document that situation to EPA and will not consider it to be impaired. COMAR §26.08.02 does include a natural conditions exception to the water quality standards.

Comment 2. Data Likely to be Unrepresentative Should Be Excluded

In some cases, the regulator may have data that is known to be unrepresentative of typical waterbody conditions. For example, water quality data gathered during a flood would not be representative of typical conditions. Where the regulator knows or has reason to believe that the data are unrepresentative, they should be excluded from use in the listing decision process. Florida excludes data collected after: (1) a rainfall in excess of a 25 year, 24-hour storm, (2) contaminant spills, or (3) discharges due to upset or bypass. Fla. Admin. Code 62-303.420(5). Florida also excludes exceedances solely attributable to NPDES violations, and "outlier values" or values due to transcription errors.

Texas addresses data typicality in part by requiring, for freshwater streams, that data be collected when stream flow conditions equal or exceed the seven-day, two-year low flow (7Q2). *Guidance for Screening and Assessing Texas Surface and Finished Drinking Water Quality*, 2000 (available on the Internet at www.tnrcc.state.tx.us/water/quality/00_305guide.pdf) ("Texas Methodology").

All of the Florida categories of unrepresentative data would be applicable to Maryland waters, and we recommend that the Methodology be amended to exclude such data. Maryland may also want to specify – based on the flow characteristics of its streams – a minimum flow sampling condition.

MDE Response: Flow regimes in the State typically vary by season, with historic patterns of "high flow" in January – June and "low flow" in July – December. The presence of sensitive life stages (larvae, sac fry, fingerlings, and juveniles) for aquatic species also correspond with high flow regimes present in the spring and early summer. Numeric water quality criteria for contaminants are normally developed using toxicity tests for sensitive life stages' acute and chronic toxicity endpoints (e.g. survival, growth, and reproduction). These criteria are a part of the water quality standard intended to protect a use designation for a waterbody. The presence of chemical contaminants at concentrations above the water quality standard may be more detrimental to the growth and propagation of aquatic life during high flow regimes when sensitive life stages are present. Therefore, it is important to perform monitoring on temporal and spatial scales that are appropriate for the monitoring objective(s).

The abnormal conditions mentioned for FL above would obviously receive the appropriate data designation ("data flagging") detailing the anthropogenic and environmental influences at the time of sampling. These types of data would not normally be used in the determination of impairments under 303(d).

Comment 3. "Hot Spot" Sampling May Be Unrepresentative

Data may also be unrepresentative of the quality of the waterbody as a whole if they are taken in known "hot spot" locations, or only in areas of the waterbody suspected to have degraded quality. As part of its data assessment, MDE should evaluate the number of samples – as well as the sampling locations – for a waterbody segment in relation to the size of the segment, the uses

of the segment, and any known sources of contaminants. In a later section of these comments, we recommend that a minimum of twenty samples be required for a listing decision. For a large or complicated segment, more than twenty samples may be necessary to characterize water quality accurately. And if all the data indicating an exceedance relate to a single sampling location, the regulator should require additional monitoring at other locations within the segment. As to repeated “hot spot” sampling, Maryland should build into its methodology a means of countering possible unrepresentativeness. For example, Florida provides that samples used to document exceedances should be “temporally independent” and requires that samples collected for chronic criteria at the same location less than seven days apart must be considered one sample, with the median value used to represent the sampling period. Fla. Admin. Code 62-303.320(4). We recommend that Maryland amend the Methodology to include a similar provision.

Maryland also should be willing to reconsider the segment size, based on the data it collects. In some instances, it might be appropriate to list only a portion of the waterbody or segment.

MDE Response: Although “hot-spot” sampling may not be representative of the waterbody as a whole, these assessments may allow for in-stream source characterization. In areas with known impairments, this type of sampling is more cost effective than a random design. For monitoring a large or complex waterbody, the sampling plan should include sufficient sample numbers to characterize the spatial extent and variation of contamination, and additional reference samples should be taken to make a valid characterization of background concentration levels, as well as sample(s) demonstrating an unimpacted or less impacted condition for that segment. MDE makes no assumptions about the spatial representation of any given sample point. Spatial impairment is determined relative to numerous factors, including:

1. Distance from outfall (point sources & stormwater – CSO/SSO)
2. Hydrologic/Geomorphic Characteristics of the segment
 - a. Physical Isolation – Dams, Check dams, waterfalls
 - b. Tidal Influences
 - c. Identification of Transport vs. Deposition Sites (sediment)
 - d. Spatial Scale - Area of integration (monitoring a 2nd order stream to integrate contributions from 2 first order streams just below confluence)
3. Sample station comparative representativeness to segment. (i.e. is sample meant to represent 100 ft, 1000 ft, 1 mile, 10 miles?) – based on history of land use in watershed – multiple sources = multiple stations?

There are numerous parameters to consider when assessing the appropriateness and applicability of chemical samples:

1. Number – in general, more is better (Real World = Budgetary Constraints)
 - a. 10% Rule - Given that you almost **never** have 10 samples for a given waterbody **period**, the time period doesn't matter. The time period that MDE is considering is over a five year? period. If you have 10 samples and you had two exceedances over a five-year period, you would list it. However, if the data are suspect or if temporal trends show decreasing concentrations, you would not list it.
1. Random vs. Targeted (PS & NPS Outfalls)

3. Ambient vs. Stormwater & Seasonal Considerations – if contaminant exceedances occur at the same time as the presence of sensitive life stages within a waterbody (e.g. spring pesticide runoff correlates with presence of fry, larvae, or juvenile aquatic life stages.

Comment 4. Thermal Exceedances Related to Variances Should Be Excluded

Under CWA § 316(a), a discharger may be granted a variance from applicable thermal WQS and limitations if it demonstrates that the standards and limitations are more stringent than necessary for the protection and propagation of a balanced, indigenous population. Wherever the state has granted a § 316(a) variance, that waterbody segment is highly unlikely to be impaired, because it supports a balanced, indigenous population. These waters should be excluded from evaluation for listing because of the unlikelihood of impairment. Alternatively, the Methodology could specify that waters with existing § 316(a) variances be deemed a low priority for impairment evaluation.

At a minimum, any exceedances related to the variance should not be used for listing evaluation. These exceedances are state-approved, and it would be inconsistent, therefore, to use them as a basis for listing.

MDE Response: The relevance of thermal exceedances to the toxics methodology being reviewed is not clear.

Comment 5. The Methodology Should Establish a Sample Number Minimum and a Statistical Approach to Determining Impairment

For listing on the basis of water column data, MDE does not require a minimum number of samples. Indeed, the Methodology would find the designated use impaired if one out of ten samples in a three-year period exceeds a single criteria. Methodology, p. 4. It is not reasonable or practical to base the listing decision on a single exceedance. A single exceedance may be an outlier value, or the result of laboratory error. And even if the datum is valid, it is inadequate for proving that an impairment exists, given the duration and frequency components of water quality criteria.

As the National Research Council stated in its recent assessment of the scientific basis for TMDLs: “[EPA’s] choice of 10 percent is meant to allow for uncertainty in the decision process. Unfortunately, simply setting an upper bound on the percentage of measurements at a station that may violate a standard provides insufficient information to properly deal with the uncertainty concerning impairment.” National Research Council, *Assessing the TMDL Approach to Water Quality Management* (2001), p. 41.

Moreover, Maryland’s approach is not consistent with what other states are doing. Florida’s new regulations are quite explicit on this point. Florida requires a minimum sample size of twenty samples, with a 10% exceedance frequency and a minimum 90% confidence level using a binomial distribution, for inclusion of a waterbody on its § 303(d) list. Fla. Admin. Code 62-303.420(2). Similarly, Texas generally requires at least nine samples over a five-year period for water chemistry, and a minimum of five samples over a five-year period for toxicity testing in water, sediment, or fish tissue. Texas Methodology, p. 4-3. Another, more rigorous statistical

approach has been developed by Dr. Robert Gibbons of the University of Illinois at Chicago. Dr. Gibbons' detailed report on his approach is appended to these comments. Dr. Gibbons' approach is appropriate for a variety of different concentration distributions (*i.e.*, normal, lognormal, and non-parametric) and offers other significant improvements over a binomial distribution approach like that which Florida has adopted.

We recommend adoption of the Gibbons approach. However, at a minimum, Maryland should follow Florida's example and require a minimum of twenty samples with a 10% exceedance frequency and a minimum 90% confidence level for listing. This statistical method is not difficult to apply. In fact, the Florida regulations contain a table showing the number of exceedances necessary for listing based on the number of samples. *See* Fla. Admin. Code 62-303.420(2) (Table 2).

MDE Response: Ideally, 20 samples would be preferable to 10 samples, but real world issues such as budgetary constraints play a major role in the extent of monitoring efforts. For this reason, professional judgment plays a key role in the determination of impacts. Some qualitative issues to consider with small sample sizes are:

1. Magnitude of exceedance: example chronic criteria chemical X = 3.4; exceedances = 3.5 and 3.6 with remainder below (<3) criteria –would not list. (If exceedance can be explained in the variability of the analytical method quantitation, would not list). If for the same example, 9 samples exist and exceedances were 5.4 and 7.6, then would list (hypothetical).
2. Correlation of sampling exceedance with seasonal presence of sensitive life stages (higher priority)
3. Number of exceedances crucial: 1 out of 6 – would not list... but how about ≥ 2 out of 4 or ≥ 3 out of 6? (Probably carries more weight than 2-3 out of 20)

We are currently assessing the incorporation of guidance (e.g. decision rules for < 10 samples; use of confidence intervals to statistical characterization of data sets) pertaining to the use of monitoring data with small sample sizes as well as the development of procedures for looking for overwhelming evidence of water quality impairment, like a single sample with well documented quality control methods shows a large exceedance of an applicable criterion (EPA – Consolidated Assessment & Listing Methodology). The Department is also looking at setting up a sampling procedure similar to and in conjunction with the Fish Tissue sampling network. Triennial sampling of individual region watersheds can be integrated with the fish tissue network. Tissue monitoring (especially accumulator species – bullhead or catfish) can be used to direct and prioritize the temporal and spatial constraints for both random and targeted water quality monitoring strategies.

Comment 6. Assuring Adequate Data Quality

For ambient water column data, the MDE plans to examine whether the data were generated by documented methods, using an “acceptable” QA/QC plan. Methodology, p. 5. With two exceptions, Constellation applauds this requirement. First, it is not enough to simply confirm that a test method has been “documented.” Unless the test method has been subjected to a rigorous interlaboratory method validation study, and its performance has been deemed to be adequately reliable for ambient testing for § 303(d) purposes, a “documented” test method is not suitable for collecting data to be used for § 303(d) purposes. The consequences of listing are far

too substantial to risk basing the process on an unreliable foundation. EPA's mandatory validation process is prescribed in 40 C.F.R. Part 136. Second, MDE should apply its data quality requirements to *all* data (not just water column data) used to make listing decisions. Any data not validated by documented methods using adequate QA/QC should be excluded from consideration.

We recommend the following minimum QA/QC criteria for data used for listing decisions:

1. All data should be collected, handled and analyzed according to fully validated and approved test methods.
2. Data that fall outside of acceptable QA/QC ranges or that are derived from tests that were not performed in strict conformance with test method protocols should be excluded.
3. Clean sampling techniques should be used when analyzing for certain contaminants (*e.g.*, mercury).
4. Data generated by an unknown source should be excluded, since its quality cannot be determined.
5. Data below the quantification or detection levels of the analytical methods used should be excluded.

MDE Response: Data used for generating the 303(d) list is received from agency sources (USGS, USFWS, USDA, DNR, MDE), or from peer-reviewed scholarly manuscripts. Additional QA info (analytical chemistry – duplicates, blanks, spikes) as well as co-occurring information (pH, hardness, AVS, TOC - when absent) is also requested from authors to insure data quality. Data from citizen monitoring groups and other “third party” sources are also scrutinized for data quality. This type of data (3rd party) can then be used to support (or refute) the primary data sources used for listing decisions.

Comment 7. Maryland Should Not List Waters Where Measures Leading to Attainment Are Planned or Underway

Even if a waterbody meets whatever rigorous approach MDE adopts for listing, as recommended in these comments, Maryland should exempt waterbodies where attainment is likely, based on planned water quality improvements. For example, where improvements to a publicly-owned treatment works are expected to eliminate a low dissolved oxygen problem within a reasonable time period, or mandated air emission requirements will substantially reduce deposition of mercury in the waterbody, Maryland should not list these waters. There is precedent for that approach. Florida provides that impaired waters will not be listed if “reasonable assurance is provided that, as a result of existing or proposed technology-based effluent limitations or other pollution control programs under local, state, or federal authority, they will attain water quality standards in the future and reasonable progress towards attainment of water quality standards will be made by the time the next 303(d) list is scheduled to be submitted to EPA.” Fla. Admin. Code 62-303.100(5).¹⁰

MDE Response: The Department would consider, on a case by case basis, waterways that meet certain attainment remediation criteria:

¹⁰ EPA's final TMDL regulations (65 Fed. Reg. 43,856 (July 13, 2000)), which are being reconsidered by the Agency, required listing of waters expected to attain water quality standards by the date of submission of the next TMDL list, but did not require development of a TMDL. 65 Fed. Reg. at 43,666 (40 C.F.R. § 130.27).

1. Measures planned: List
2. Funding Approved and Implementation Started: List, but prioritize for removal based on results of future monitoring efforts.
3. Project Implementation complete and early sampling indicates remediation target will be met. Consider removal, contingent on continued improvement based on subsequent monitoring protocol within a given time period (e.g. 3-5 yrs).

Comment 8. Listing on the Basis of Sediment Contamination

In general, Constellation agrees with Maryland's decision to use a weight-of-evidence approach for determining sediment impairments. Table 1 (Possible Conclusions Provided by Using the Sediment Quality Triad Approach) explains how the weight-of-evidence approach will work, and the judgments reflected in the table (e.g., do not list where sediment toxicity tests are negative, but there is biological community alteration and sediment chemical contaminant tests are positive) are appropriate.

However, Scenario 6 of Table 1 requires some clarification. In that scenario, tests of sediment toxicity and sediment chemical contamination are positive, but there is no alteration of the biological community. The table states that, in this scenario, the site will not be listed, but additional monitoring will be required. The description of Scenario 6, however (Methodology, p. 11), provides that if there is no biological community data or the available biological indicators are not applicable to the waterbody in question, MDE "will apply its best professional judgment, but will likely determine that the designated use is not supported." There are several problems with this approach. First, best professional judgment ("BPJ") may be appropriate to compare the weight of apparently different outcomes, but it should not be used to fill a data gap as significant as biocriteria. That is particularly true given the severe limitations associated with bulk chemical and toxicity testing of sediment. Second, in many cases there will be some form of information on the status of the biological community in the area. While the data may not be sufficient under MDE's biocriteria program, it may be sufficient to determine that there is little likelihood of impact to the community. In this instance, the water should not be listed. Third, listing of the waterbody may waste resources, because it may be determined through later monitoring that the biological community is in good health. Since the health of the community is the overall objective, it makes sense to acquire this information before rendering a listing decision.

Therefore, we recommend that MDE revise its description of Scenario 6 to require additional biological monitoring, rather than listing, in the absence of sufficient biological data to render a listing decision. We also recommend that MDE consider widening what it considers to be adequate biological data for purposes of listing decisions. For example, MDE may want to consider state fisheries assessment data.

MDE Response: In Scenario 6, best professional judgment will only be used when the benthic community component cannot be met (no biological community data or the available biological indicators are not applicable to the waterbody in question). Best professional judgment includes:

1. Comparison of chemical concentrations to established sediment quality guidelines (ERM (ERM-Q), PEL) – measure magnitude of exceedance (level = ERM ≠ level = 5 ERMs) and apply

modification factors if appropriate (AVS-SEM ratio, calculation of anthropogenic enrichment (metals); TOC (organics))

2. Compare toxicity relative to absolute survival in controls and weight tests accordingly (e.g. if control survival was 91%, mortality of 90 % vs. 50 % vs. 15 % would be weighted differently), assuming impact of all confounders were minimized.

3. Statistical comparisons (e.g. via linear regression) of measured chemical concentrations vs. laboratory toxicity tests.

4. Compare benthic alteration test results from areas impaired by similar levels of toxic contaminants, and similar observed laboratory toxicity, where available. (e.g. Baltimore Harbor and Back River).

The Department will clarify Scenario 6 in its final methodology release to read “Exceedance threshold-based listing – See Comments for explanation”, with 1-4 above in comments section.

The temporal relationship between the triad components is crucial, so fisheries data may be unsuitable. Does fisheries perform benthic sampling (B-IBI)? – presence of fish in an area may give very little information relating to the community structure in the sediment. In addition, the fish IBI only validated in MD freshwater systems.

Comment 9. Listing on the Basis of Fish Tissue Data

MDE proposes to list waters as impaired if a fish consumption advisory has been issued.

Following EPA guidance, MDE notes two exceptions to this general listing requirement:

(1) when the state has issued a statewide or regional fish consumption warning, based on data from a subset of waterbodies and species, and (2) when the state has set a high consumption value to protect a specific sensitive population. Methodology, pp. 12-13.

Even when dealing with a waterbody-specific fish consumption advisory, it is not clear that listing is necessarily appropriate. Migratory fish present in the waterbody on a seasonal basis may have ingested the contaminant of concern largely in another locale. This is not to suggest that states should not use fish advisories. So long as a state has developed fish tissue levels based on an adequate scientific process and in accordance with state administrative procedural requirements, it is appropriate to use those levels to inform their listing decisions. But all information based on such advisories should consider whether the fish tissue levels observed are attributable to conditions in the waterbody being considered for listing.

MDE Response: The Department’s Fish Tissue Monitoring Program collects samples from a variety of species, including resident commercial, game, accumulator, and nongame fish and shellfish species, and seasonally resident commercial, game and nongame fish and shellfish species. There are two conditions in which fish tissue data would be used for listing on 303(d):

1. Any General Population consumption advisory due to elevated contaminant levels in resident accumulator species (e.g. bullhead, catfish, clams, oysters).
2. Any General Population consumption advisory due to elevated contaminant levels in any species from a lake or impoundment (species residing in these waters are by definition resident).

Comment 10. Blanket Adoption of EPA § 305(b) Guidance is Unwarranted

MDE states that in general it has adopted EPA's § 305(b) guidance for identifying waterbody segments impaired due to chemical contaminants, but adds that "there may be instances where our determinations may vary based on scientifically defensible decisions." (Methodology, p. 3). As already noted, MDE should abandon EPA's 10% exceedance rule for determining impairment. That approach is far too simplistic for making decisions bearing the TMDL consequences associated with the § 303(d) listing process. Use of the Gibbons approach or the Florida approach would greatly enhance the reliability of Maryland's listing decisions. Therefore, Maryland should not generally endorse EPA's § 305(b) guidance. Instead, it should strive to ensure that all aspects of its listing methodology are designed to identify impairments based on sound data and information. In this way, Maryland will produce a strong methodology that will serve its TMDL program well. Additionally, EPA is expected to issue listing guidance within the next few weeks. We urge MDE to consider that guidance and open up a brief comment period to seek public input on possible changes to the proposed MDE listing methodology.

MDE Response: The Department agrees that blanket adoption of impairments as indicated by CWA 305(b) is not appropriate, since the impairment decision under 303(d) must be scientifically robust, since these are the basis for the TMDL process. While the Department agrees that EPA's "10% rule" (>10%) may be too simplistic as the basis for setting TMDL load allocations, it is and has been a widely used "rule of thumb" for making designated use-based impairment decisions under CWA 303(d) for many states and tribes. Setting TMDL load allocation for Maryland involves a comprehensive assessment of all point and nonpoint sources affecting a waterbody. MDE will continue to review all proposed guidance (e.g. CALM), and will consider modifying the proposed listing methodology based on technical advantages, consistency, and budgetary constraints of the affected agencies.

Comment 11. EPA does not provide guidance for interpreting water quality data for the purposes of developing the 303(d) list. However, EPA does provide guidance on making use support determinations for the State Water Quality Assessments (305(b) Report) (EPA 1997). In general, MDE adopted the 305(b) guidance for identifying waterbody segments impaired due to chemical contaminants. Even though the Department will adhere to these methods as closely as possible, there may be instances where our determinations may vary based on scientifically defensible decisions. EPA issued a Stressor Identification Guidance on 12/00 that may offer help with this problem. It is important to note that there maybe situations which do not support an impairment determination from chemical contaminants, but rather from another stressor (e.g. dissolved oxygen, biocriteria), and would therefore be addressed elsewhere. This document provides the specific methodology used by MDE for identifying waterbody segments impaired due to *chemical contaminants*.

MDE Response: Incorporate the reference. (EPA 1997). In addition, the EPA recently issued a Stressor Identification Guidance Document (EPA/822/B-00/025; December 2000) that aids in the identification, analysis, and characterization of potential causal stressor sources. Generally, this document refers to stressors in the context of biocriteria development. The Department will also incorporate guidance from this document where specific reference is given to chemical stressors.

Comment 12. Seasonality of behavior or presence of AVS or ammonia and pH further complicate this issue and have to be taken into account when scheduling sampling period and sampling protocol and comparing bioassay results with sediment chemistry and community structure. This spatial and temporal behavior of the sediment component must be understood when evaluating test results.

MDE Response: The Department agrees, however, this understanding does not always yield conclusive results, thus reinforcing the need for other supporting data. Seasonality of behavior in regards to benthic community structure is considered in the scheduling of sampling protocols for Benthic IBIs. Analysis of the community structure when combined with appropriate sediment chemistry can aid in the elucidation of in situ conditions (e.g., AVS) that may require adjustments in the laboratory to control for superfluous or artifactual adverse effects when conducting toxicity tests. Bioassays using field-collected ambient sediment samples and laboratory-reared or field-collected organisms attempt to evaluate the toxic nature of the sediment while controlling for endemic geochemical conditions that may act as toxicants to the test organisms under laboratory conditions.

Comment 13. (EPA) The five EPA SQC are about as final as we are going to get and should be used as such along with EPA ESG s.

MDE Response: The Department relies on the EPA SQCs/ESGs as guidance combined with other sediment screening metrics (e.g. MDE SQBs, NOAA ER-Ms) and other site-specific information (e.g. AVS content, TOC) when deriving a numeric criterion to evaluate sediment chemistry as a component of the impairment decision-making process. Note: The Department considers the EPA SQC/ESG criterion as guidelines only, since they could not withstand the technical regulatory scrutiny required for adoption as final criteria. Both the quality of sediment chemistry data and associated screening thresholds are considered when conducting an evaluation. Once the quality of data has been established, the potential for adverse effect from contaminated sediment is said to be high (Define what is meant by high.) if either of the following conditions are met.

MDE Response: *Change high to probable.*

Comment 14. Furthermore, various environmental conditions in the sediment can have a profound effect on the availability and toxicity of the sediments to aquatic environment (e.g., AVS for metals, organic carbon for organics, etc.). If data on these parameters are available, MDE will use best professional judgment to interpret the effects of these parameters on the sediment chemistry data. AVS and organic carbon can cause shift in effects to the point that a magnitude change in concentration can have no appreciable effect on toxicity.

MDE Response: *Change to ..."interpret the effects of these parameters on the sediment chemistry and apparent toxicity potential on a site-specific basis."*

Comment 15. The mere exceedance(s) of a sediment threshold, however, does not in itself establish an adverse effect from toxicity, but helps to identify the chemical that might be

responsible for any observed adverse effects from toxicity. Given these limitations, MDE does not believe that the exceedance(s) of sediment thresholds are appropriate as sole indicators of use attainment. Magnitude of exceedances as well as presence of AVS, organic carbon, or pH levels does play important role. Instead, we recommend using all three data components as a basis for interpreting narrative criteria and developing pollutant reduction strategies. Does this include the five pesticides for which both SQCs and ESGs have been developed.

MDE Response: Add: "Although the magnitude of exceedance of a screening value can indicate a potential for toxicity, the presence of AVS, TOC, or unfavorable pH conditions play an important role in the bioavailability of contaminants." 2nd Question – The Department advocates the use of the Sediment Quality Triad (SQT) approach inclusive of the EPA SQCs and ESGs for pesticides due to the fact that these guidance numbers have not been published as “Final Criteria” by EPA. This weight of evidence approach provides a more thorough analysis of site data for listing possible impairments, than a sediment screening value alone."

Comment 16. Can marine results/methodologies be readily applied to freshwater conditions?

MDE Response: The Anderson reference only attempts to illustrate the correlation between toxicity observed in the laboratory and benthic community impairment in situ. Adoption of methodologies and comparative data usage may be appropriate in the estuarine portions of Bay tributaries and the mainstem if similar environmental conditions exist. The Department does not advocate the application of marine methodologies to freshwater systems. Extensive freshwater validation of marine test methodologies would be required prior to the consideration of adoption of the aforementioned techniques.

Comment 17. Extreme values in one component may be sufficient to make impairment determination. (i.e., EPA's Policy of independent applicability).

MDE Response: Given sufficient data quality, extreme values in one component should trigger the timely monitoring of the other triad components at the site so that a weight of evidence determination can be for any pending impairment. Basing an impairment on a single extreme value, especially when using data or methodology that is not recent may lead to an improper determination of impairment, particularly if other factors (AVS, TOC for chemistry; seasonality or predation for BIBI, or uncontrolled ammonia or D.O. confounding in toxicity testing) were not accounted for. When presented with extreme, high quality data from one triad component, a timely triad-based assessment of the listed site, followed by an update of the impairment listing would result.

Comment 18. The decisions made under scenario 3 of table 3 are O.K. with some potential exceptions. Biological magnification potential may be affected by seasonality of AVS and organic carbon.

MDE Response: Scenario #3 makes the assumption that seasonality of confounders (e.g. AVS) was an a priori assumption of the study design and is accounted for.

Comment 19. We disagree with the listing decision made in scenario 4 of table 3. Bioassay tests are normally run on multiple runs with different critter types and dilutions Independent applicability may be used to make decision of impairment.

MDE Response: From Slide 1 of the Pendergast briefing, the Independent applicability Rule makes the assumption that multiple types (chemistry, toxicity, biological community) of data are present, and that a single type or source indicates an adverse effect. Please refer to Table 1. for decision scenarios for multiple data sources (e.g. SQT). Normal protocol for ambient sediment toxicity testing is one organism, with no dilutions. If the magnitude of toxicological effect was great on several runs with multiple test species, the Department would then proceed with triad-based monitoring and re-evaluation of the listing decision based on best professional judgment.

Comment 20. We disagree with the listing decision made under scenario 6 of table 3. Under MBSS decision processes, the Benthic IBI can take precedence over Fish IBI when certain conditions prevail. What's the protocol when B-IBI and F-IBI are in conflict? However, this biological data is mostly used for watershed level assessment determinations and may not be applicable in reaches of streams at the level of detail that a toxicity determination is made.

MDE Response: The Department does not understand how the aforementioned comment supports EPA's disagreement with the scenario #6 justification. EPA states that the B-IBI can take precedence over the F-IBI under certain conditions. The Department understands the context of use of the B-IBI for watershed level assessments, but when the B-IBI is used as a component of the SQT, it is a good indicator of use attainment. The B-IBI can also reveal a parameters' effects such as AVS concentrations leading to toxicity mitigation in situ, but the same AVS leading to toxicity in ambient sediment toxicity testing in the laboratory. Additional monitoring would be warranted at this location due to the indication of a "stressed" system as evidenced by observations of toxicity and elevated chemical concentrations in the sediment.

Comment 21. This statement appears to conflict with earlier statements regarding decision making in Table 1. A clarification may be needed. Maybe it should be explained that single indicator is not overwhelming if seasonality of AVS, OC, or pH is noted

MDE Response: This statement makes the assumption that several data sources exist, and that the indicators' significance is overwhelming. Some examples may be:

- a. Sediment concentrations that are 25-100X the ER-M
- b. Toxicity tests with 75-100% mortality in the exposed group vs. < 10% mortality in control groups.
- c. Total absence of benthic infauna from several grab samples at a sampling location.

Comment 22. This statement needs to be reconsidered in regard to the following condition: large magnitude of exceedance with regard to four conditions listed in Sediment Chemistry Data section. EPA's Policy of Independent Application may apply.

MDE Response: Condition d) of the aforementioned listed conditions mentioned sediment toxicity data as one of four conditions. The Department states (2 paragraphs previous to this

one) “However, if data are not available for all three components, the Department will use its discretion but will consider an impairment determination if b) If a toxicity test shows toxicity and is confirmed either by chemistry data or a degraded biological community, its designated use is not likely supported and an impairment determination will likely be concluded”. The Department agrees with the EPA statement when taken in the same context as the previous MDE Response from the previous comment.

Comment 23. 303(d) –Fish Advisories If this specific sensitive population advisory was generated from actual data in specific watersheds, then those waters must be listed as impaired, even if that means that only part of a regional advisory waters are listed.

MDE Response: The Department feels that this comment is contradictory to guidance from the EPA Office of Water guidance (WQSP-00-03) received by the Department 24 October 2000. Chart 1 (page 6) states that if the risk assessment parameters (e.g. consumption rate, body size, exposure duration) are more protective (as would be indicative of a sensitive population advisory) than the criteria specified in State or tribal water quality standards, the Department would list as threatened {305(b)}, rather than 303(d) due to uncertainty regarding a water quality exceedance.

2. Chemical Contaminants: Page 4, bottom of page: If there are less than 10 samples for a specified waterbody, MDE will interpret available data on a case-by-case basis... In the case that no criteria are available for a particular contaminant or no criteria are exceeded, other impairment indicators (e.g., ambient water column toxicity data) will be evaluated using best professional judgment. During this evaluation process, if toxicity is indicated, a Toxicity Identification Evaluation (TIE) may be considered to further identify the possible contaminant sources(s) causing toxicity... When warranted, MDE will also utilize spatial and temporal trend analyses as an additional evaluation tool for making impairment determinations.

Comment 24: The determination that ... is subjective, especially since these are cases where “less than 10 samples for a specified waterbody are available” and “no criteria are available or no criteria are exceeded”. Criteria should be made for contaminants as a basis of the judgment

MDE Response: The Department recognizes the need for criteria regarding sample number, spatial representation, etc., and is currently working on developing statistical methods for determining appropriate minimum sample numbers based on basin segment size. Regardless of the determined methodologies, there are cases when the magnitude and or frequency of a WQ criteria exceedance warrants an impairment determination, even when there is a small # of samples. An example might be 4/6 ambient water samples exceeding chronic criteria for pesticide X. Where only 1 sample exceeds criteria in a sample size < 10, several qualifiers must be examined prior to considering basin segment for impairment.

1. Magnitude of exceedance (acute vs. chronic/for consideration of chronic exceedance, should be well above threshold; e.g. > 1.5X)
2. Data reliability – QA/QC
3. Temporal – ambient vs. stormwater; ambient high vs. ambient low

4. Other indicators of impairment – weight of evidence (sediment chemistry, SQT, fish tissue concentrations).

Comment 25. Toxicity Reduction Evaluation (TRE) and their associated TIE procedures were developed for industrial and municipal point source discharges where contaminants in the waste stream do not vary significantly from day to day, and can be systematically identified and traced to their source. TRE and TIE procedures are not as useful for identifying toxicants in storm water discharges due to the intermittency and variability of storm events and their resulting runoff contaminant levels. For this reason, it is doubtful that TRE and TIE procedures would be successful in identifying toxicants in ambient waterbody samples where point source and non-point source inputs vary, and where other ambient factors in the water column and sediment contribute to constituent variability. We therefore recommend that results from TIEs not be used as a sole basis for listing a waterbody.

MDE Response: The Department recognizes the shortcomings of TIE methodology for nonpoint stormwater discharges, however these techniques may help to identify a previously missed stormwater component that may have caused toxicity under ambient conditions. In any case, additional follow-up sampling would be required to verify the frequency, possible sources, and toxic potential of the identified constituent prior to listing under 303(d).

The Department agrees that spatial and temporal trend analyses should be used as an additional evaluation tool for making impairment determinations.

Comment 26: There is no disagreement that sediments are an integral component of aquatic ecosystems. However, we disagree with the Department's proposed approach for using and interpreting sediment data in making impairment decisions. It appears from the draft decision criteria that a waterbody could be listed as impaired based on sediment data alone, even if there are no violations of water quality standards in the water column. If that is the Department's intent, we strongly disagree. Sediments, for which no numeric standards have been approved, cannot outweigh approved water quality standards in making impairment decisions.

MDE Response: Example: High PCBs > (ERM) in the sediment contribute to food-chain based bioaccumulation of PCBs in edible finfish, even though no WQ criteria (water column) for PCBs is exceeded. A "general population" fish consumption advisory results, and water is listed due to non-support of designated uses (fishable/swimmable) – (EPA Doc WQSP –00-03). The listing of a waterbody based solely on sediment will be dependent on several parameters.

1. Number and temporal/spatial extent of samples.
2. Number of contaminants exceeding SQG's
3. Magnitude of exceedances
4. Other triad data for sites (B-IBI and Tox testing)
5. Co-occurring data (AVS, TOC)

Specific comments are as follows.

Comment 27. In the first paragraph on page 4 of the draft criteria document is stated that, “the Department considers all existing readily available...data” in determining if a waterbody segment should be listed. This should only apply to data of sufficient quality.

MDE Response: The Department agrees with WSSC that impairment decisions should not be made without first determining the veracity of the data through technical scrutiny. The data quality requirements set forth in most current scientific research and regulatory programs ensure a certain “minimum” level of data quality. Data quality is also assured through the peer review process, as well as the QA/QC process and other GLP requirements. Data collected through in-house contractors are also scrutinized by technical experts prior to their use. If the data quality is unacceptable, but indicate an impairment, the segment is targeted for additional monitoring. In general, EPA’s report guidance (1997a) utilizing a two-layer approach to assessing data quality is followed.

Comment 28. The draft criteria document states that, “In the case that no criteria are available for a particular contaminant or no criteria are exceeded, other impairment indicators (e.g., ambient water column toxicity data) will be evaluated using best professional judgment.” This implies that the Department may list a water as impaired even if water quality criteria are met, or for impairment caused by some constituent for which a water quality standard has not been developed. We strongly disagree with either of these approaches. We question how an impairment decision can be made and a TMDL developed for a constituent for which no water quality criteria exist. If the Department intends to assess toxicity based on constituents for which there are no existing water quality criteria, such criteria should be developed and adopted as water quality standards before an impairment determination is made.

MDE Response: This situation is the reason for the narrative water quality criteria that set goals for water quality and allow a determination of use attainment to be made when numeric water quality criteria are not available. If a waterbody were to demonstrate acute toxicity because of a substance(s) for which a water quality criterion did not exist, it would not be appropriate for the Department to ignore the toxicity simply because of the lack of the criterion.

Comment 29. The draft as written appears to indicate that a single sample could cause a waterbody to be listed. It erroneously equates “>10%” with “1 out of 10”. Also, if there were 9 samples at 1% of the WQS and one sample at 101% of the WQS, the water could be listed. Due to sampling and analysis variability alone, in no case should a body of water be listed on the basis of a single sample. At a minimum, 2 out of 8 samples should indicate substantial impairment.

MDE Response: The 2000 Maryland 305(b) – Appendix E-5 report states that “if more than 10% of the observed water quality measurements at a site exceed the defined water quality criteria for an applicable use, the waters may be considered impaired.” If only 10 samples were taken, then at least 2 would have to exceed WQS. Other considerations in determining impairment are temporal sampling regime (presence of sensitive life stages?); magnitude of exceedance; exceedance of chronic vs. acute WQ criteria; and ambient vs. stormwater sampling. For example, an ambient sample exceeding acute criteria would carry more “weight” than a stormwater sample exceeding chronic criteria.

Comment 30. Many people involved in sediment criteria development concur that sediment criteria do not have sufficient scientific basis to be adopted as water quality standards. By this document, the Department is attempting to adopt sediment criteria as standards. We do not believe that the Department should attempt to derive sediment quality benchmarks (SQBs) if the USEPA has not been able to do so.

MDE Response: The Department is well aware that the scientific community does not agree on a single number with which to set numeric criteria for sediment contaminants. MDE is also aware that there is scientific consensus that a “weight of evidence” approach using the “Sediment Quality Triad” may be the best method to deal with the uncertainties inherent in a sediment impairment evaluation. MDE has the responsibility to provide for the protection of waterways for aquatic resources and recreation, where attainable. Since the sediment is an integral component of the aquatic ecosystem, sediment quality can have a direct effect on the attainment of a designated use. Consequently, it is necessary and appropriate for the Department to assess and protect sediment quality. The Department will first make a narrative determination of impairment, if sediment quality is the only component with an observed impairment. Furthermore, the Department will use the “Sediment Quality Triad” approach as a tool to aid in the quantitative characterization and interpretation of the aforementioned narrative criteria statements, thus enabling the determination of water quality standards’ violations from contaminated sediments.

Comment 31. On page 7 of the draft criteria document is a statement that reads, “Once the quality of the data has been established, the potential for adverse effect from contaminated sediment is said to be high if either of the following conditions are met...”. We agree that there is a potential for adverse effect, but do not agree that it is necessarily “high”.

MDE Response: The Department has received several comments about the term “high” and will replace with the word “probable”. The connotations of the word “high” are easily misinterpreted.

Comment 32. We strongly disagree with the proposed approaches for interpreting sediment data as outlined in the four conditions on page 7 of the draft criteria document. Authors of ER-Ms have described ER-Ms as being non-regulatory, yet the Department proposes to use them for regulatory decisions. The Department’s proposed approach would result in a finding of a high potential for impairment based on a single comparison of sediment chemistry data to an ESG or ER-M, even though USEPA in their National Sediment Quality Inventory required that sediment chemistry data exceed two separate benchmarks for a finding of probable adverse effects, due to the limitations of ER-Ms. Furthermore, the Department does not describe how sediment chemistry data will be interpreted for comparison to ESGs or ER-Ms: will a single data value exceeding the ESG or ER-M comparison criteria be sufficient cause for a finding of probable adverse impacts? We believe that the Department needs to present a more rigorous evaluation and supporting information on the use of sediment chemistry data if sediment chemistry comparison procedures are to be used in making listing decisions.

MDE Response: The Department recognizes the shortcomings of published numeric screening guidelines for the assessment of sediment toxicity. The Department is not using these thresholds in a regulatory sense as criteria, but does appropriately use them as assessment tools. Further, MDE does not advocate the use of a single benchmark exceedance for an impairment determination, nor does it advocate the use of sediment chemistry as a stand-alone guideline upon which to base an impairment listing decision. The sediment chemistry screening value should be used in the context of the “Sediment Quality Triad” assessment of contaminated sediments. If there is a large magnitude exceedance of an SQG, additional monitoring to include toxicity testing and benthic community indices should be conducted at the site. If adverse effects are observed in all three data components, an impairment listing should be made. A toxics impairment listing can also result if there is evidence of toxicity and community alteration, but no supporting chemistry. This scenario is indicative of a toxic substance with no threshold, or some unmeasured contaminant. Additional monitoring is required for identification of the substance. A possible solution to the lack of standardization of the SQG component is the use of a “Consensus-Based” maximum concentration guideline (Probable Effect Concentration-PEC), an approach proposed by MacDonald et. al 2000). The published PECs (derived from at least 3 separate SQGs meeting acceptance criteria) were able to predict the toxicity of 28 common sediment contaminants (8 metals, 10 PAHs, Total PCBs, and 9 OC Pesticides) found in freshwater ecosystems. The PECs were able to correctly predict the toxicity of at least 20 sediment samples with 75% (minimum) accuracy (chlordan, 73, and endrin, 37.5% were exceptions). In addition, combined PEC quotients were calculated to predict the toxicity of chemical mixtures in sediments. Of the total 174 samples with PEC Quotients < 0.5, 83% were correctly predicted to be non-toxic. Conversely, 90.3% of the total 441 samples with PEC Quotients >0.5 were correctly predicted to be toxic. Similar derivations may be possible for contaminants in marine and estuarine systems.

Comment 33. The Department rightly notes that sediment and other bioassays can be inconsistent and frequently produce false positives. A study by the Western Coalition of Arid States showed a significant number of false positives for the Whole Effluent Toxicity test when using sample blanks. We also agree with the conclusion that toxicity in a sediment bioassay under laboratory conditions does not mean that sediments are toxic under in-situ conditions. The Department notes that bioassays alone should not be used as the basis for making listing determinations. The demonstrated problems of false positives and toxicity under laboratory conditions with sediment and other bioassay tests clearly shows that the sediment bioassay is not a reliable measure of in-situ sediment toxicity. Therefore, the use of sediment bioassays should not be used in assessing potential toxicity and in making impairment decisions.

MDE Response: The Department recognizes the shortcomings of sediment bioassays, but the existing ambient toxicity testing methods have been refined so that the confounding effects of in-situ vs. laboratory environments are minimized. The Department once again does not advocate the use of a sediment bioassay as a sole indicator in an impairment listing decision, but rather as a component of the sediment quality triad approach. A false positive in a sediment bioassay using the Triad approach would easily be “teased-out”.

Comment 34. The Department presents in Table 1 eight scenarios of various combinations of results of sediment toxicity bioassays, sediment chemistry data and benthic community surveys

and possible listing decisions. Also provided is discussion of the interpretation of results and likely listing decisions. the Department's proposed "weight of evidence approach" for interpreting results under the Sediment Quality Triad is inconsistent.

A "+" in the table indicates a "measured difference" between test and control or reference condition. Yet the draft nowhere defines what constitutes a "measurable difference."

MDE Response: The Department will clarify what a measurable difference is for each of the following Triad components:

1. Toxicity testing: statistical significant difference as compared with control or reference site using an accepted test (e.g. T test, ANOVA, linear regression)
2. Sediment Chemistry: any exceedance of an SQG (individual or Consensus-Based)
3. Benthic IBI: numeric index score is compared to a desired guideline (reference waterway condition or restoration goal) and degree of impairment is determined. The B-IBI uses accepted population metrics such as Shannon Diversity Index, and species abundance and biomass of pollution-tolerant and pollution-sensitive species. The B-IBI developed for the Chesapeake Bay is score from 0-5 with values lower than 2.7 considered degraded.

Comment 35. Scenario 6 (positive bioassay, positive chemical and negative benthic community alteration) would not result in a listing for toxics. We agree with this conclusion. But on page 12 of the draft criteria it is stated that where no biological data exist or where those indicators are not applicable, the Department will base impairment decisions on "best professional judgment, but will likely determine that the designated use is not supported". This statement implies that the Department will make listing decisions based only on sediment toxicity and chemical data, even if benthic community alteration has not been demonstrated, which contradicts the intent and very foundation of the sediment quality triad approach.

MDE Response: The Department will use its best professional judgment when faced with the above situation. If benthic community data does not exist or is not available, the impairment listing decision will be based on the magnitude of exceedance for the remaining two components. An example: Sediment Chemistry: # of samples exceeding guideline (e.g. 5/6 exceeded ER-M for Ni); or magnitude of exceedance of single sample (5-10X ER-M); or 1 sample with exceedances of several contaminant SQGs (e.g. > ER-M for Cd, Ni, and PCBs) Toxicity Testing: Less than 50% survival vs. 90-100% for controls (several replicates or different species), or 100% mortality vs. > 80-90% survival for controls (single species).

Comment 36. Scenario 7, with a positive bioassay, a negative chemical test and an altered benthic community, would result in a listing for toxics. We disagree with this conclusion because the same results could derive from a false positive bioassay and alteration of the benthic community due to some other cause, such as habitat or temperature modification. A similar argument could be made for Scenario 1. Therefore, unless those causes of alteration of the benthic community not related to water quality can be explicitly ruled out, the use of the sediment quality triad as proposed is not a reliable measure of impairment.

MDE Response: The Department acknowledges the possibility of confounding in the assessment of benthic community alteration. When good quality data suggests no confounding factors, and toxicity is present, then the waterbody may be impaired due to the additive or synergistic actions of two or more chemicals present at concentrations below their respective water quality criteria or sediment guidelines. Magnitude of effect for alteration and toxicity is also a criteria for this decision process. Another possibility is the presence of a chemical for which no criteria exists, in which case a TIE would follow to identify the chemical in question. The impairment listing decision is appropriate because there is a link between toxicity and community alteration. The TMDL could not begin until the chemical is identified and a numeric criterion adopted.

**Stratified Lakes/Dissolved Oxygen
Interpretation of Dissolved Oxygen Standards in Maryland's
Thermally Stratified Lakes**

List of Commentators

Author	Affiliation	Date
Patricia Gleason and Elaine Harbold	U.S. EPA-Region III	6/15/01
Cy Jones and J.L Hearn	Washington Suburban Sanitary Commission	6/18/01
Steve Stewart	Baltimore County DEPRM	6/21/01
James M. Stuhltrager	Mid-Atlantic Environmental Law Center	7/31/01

Comments and Responses

Comment 1. Nutrients, clarity and chlorophyll levels are normally used in determining the trophic state of lakes.

MDE Response: MDE considers these criteria, as well as the professional judgment of our sister agency (DNR) and other lake authorities.

The natural evolution of lakes is toward eutrophication (Reid 1961). Eventually, ecological succession by marsh, meadow and forest follows, unless human intervention slows or reverses the process.

Comment 2. Does this process really apply to a large river valley reservoir.

MDE Response: Yes, and probably more so in a typical reservoir system, since the ratio of watershed area to impoundment surface is usually larger than in a natural lake system. This engenders greater nutrient and sediment loading.

Often, stratified lakes do not exhibit a separation into three distinct layers. The epilimnion is typically present as defined above; however, temperature in the underlying waters may decrease continuously down to the lake bottom. In this document, the term “hypolimnion” is used to define waters below the epilimnion, regardless of whether the lake exhibits three-layered thermal stratification.

Comment 3. If a lake does not exhibit the properties of stratification with a metalimnion layer where there is a 1 degree C. change per meter, they that lake will have to be considered completely mixed and the minimal 5 mg/l DO criteria has to be used throughout the water column for determining use attainment.

MDE Response: The temporal and spatial occurrence of stratification are not static. Occasionally, available data may not be sufficient to identify and delineate an obvious limnetic zone. In such cases, if there is a reasonable expectation of the occurrence of thermal stratification, managerial decisions may require the collection of more data or the application of best professional judgment.

Additionally, many lakes in Maryland are completely mixed, even in summer. In these cases, the existing surface water quality standards as enumerated in COMAR are considered to apply.

Comment 4. In regard to lakes (such as Deep Creek Lake) where stratification may only occurs briefly in the many branches, will the assessment be conducted using the epi- or hypolimnion criteria.

MDE Response: The epilimnion is by definition surface water; thus, the existing surface water quality standards as enumerated in COMAR are considered to apply.

At times of stratification, the interim interpretation applies to the hypolimnion only. Further refinements of the interim interpretation may be incorporated as needed to address instances where portions of a waterbody undergo differing degrees of stratification.

Comment 5. How were the interim trophic states determined (Carlson’s TSI or other method)?

MDE Response: The term “interim” applies to MDE’s interpretation of the DO standard as applied to stratified lakes, not to the trophic state of the lake or impoundment.

Comment 6. The troubling aspect of this procedure is that Maryland accept DO levels at and below the hypoxic level of 2.0 mg/l as within its fully supporting designated use category even though this levels are both chronically and acutely toxic to fish and other organism. No reference is made to the extent of this violation in the water column. Will a lake with only a few

meters (or small % of water column) of bottom waters below the minimal hypolimnion % saturation be looked upon in the same light as a lake with a high % of bottom waters below the minimal hypolimnion % saturation level? The magnitude of the violation will be different and will affect a larger % of the lake volume or acreage. Lastly, the issue of persistent (10 - 50 days) vs. episodic (0 - 12 days) of low DO saturation was not touched upon. In the end, Maryland's criteria for DO does not allow for these exceptions. Maryland may have to do for their lakes what is done in the Chesapeake Bay and characterize what part of the total lake acreage fails the 5.0 mg/l DO criteria when listing. The listing category for 303b purposes will be determined by the available information.

MDE Response:

- Maryland's interim interpretation acknowledges the naturally-occurring phenomenon of hypolimnetic hypoxia in some stratified lakes.
- As refinements to the interim interpretation are incorporated, Maryland will consider the commentators' suggestion of considering hypoxic volume-days in the overall analysis.
- Maryland, like all States, has a limited budget to allocate to data collection. Hypolimnetic water quality conditions cannot realistically be monitored on a continual basis. Maryland's interim interpretation provides a mechanism to address hypolimnetic DO under critical conditions (i.e., maximum seasonal thermal stratification).
- During the current Triennial Review period, Maryland will consider revising the application of the DO criterion to stratified lakes.

Comment 7. This procedure also provides a reasonable approach for TMDL listing. In some cases, particularly for drinking water reservoirs, where eutrophication can cause an increase in water treatment costs. There are efforts under way to change the trophic status in some of these situations. Footnote 2 indicates that the selected trophic status is compatible with the lake's designated use. There should be more emphasis on this.

MDE Response: Maryland appreciates the commentator's desire to maintain flexibility in the selection of an impoundment's trophic status. Assignment of trophic status is dependent upon professional judgment. There is no universal procedure for determining an impoundment's desired trophic state. This is one reason for keeping this decision methodology in guideline—as opposed to regulation—form. The determination of an impoundment's desired trophic state is a policy decision. As an interim tool for impairment decision and TMDL development, Maryland has been using trophic state as described in the 305(b) reports. A final determination of an impoundment's desired trophic state is made when additional monitoring is conducted prior to the development of a TMDL.

Comment 8. While its proposed approach seems logical and reasonable, the draft decision document stops short of identifying specific listing criteria. However, as with the other guidance numeric criteria discussed here, the Department is proposing to use guidance values as water quality standards.

MDE Response: See response from OAG to this comment in other sections.

Comment 9. Our evaluation of the proposed interpretation does not include an evaluation of the trophic status designations, which were referenced but not included in this draft document. Application of the procedures assumes that a reasonable and appropriate trophic status is selected for the lake in question. However, we note that while exceedances of the minimum hypolimnetic dissolved oxygen saturation could indicate impairment, they may also indicate that the assigned trophic status may not be appropriate. We suggest that when exceedances of the hypolimnetic standard are found, a review of the trophic level *classification* should also be performed before the decision is made to list the lake as impaired.

MDE Response:

- Assignment of trophic status is dependent upon professional judgment. There is no universal procedure for determining an impoundment's desired trophic state. This is one reason for keeping this decision methodology in guideline—as opposed to regulation—form. A final determination of an impoundment's desired trophic state is made when additional monitoring is conducted prior to the development of a TMDL.
- The determination of an impoundment's desired trophic state is a policy decision. As an interim tool for impairment decision and TMDL development, Maryland has been using trophic state as described in the 305(b) reports.

Comment 10. The assignment of “Minimum Hypolimnetic Dissolved Oxygen Saturation” to the meso-eutrophic, mesotrophic and oligo-mesotrophic trophic levels based on equal increments of 23% shown in Table 2 of the draft criteria seems somewhat arbitrary. We suggest that a minimum hypolimnetic DO saturation of 10% be applied to all classifications within the mesotrophic range until a more scientific basis for appropriate values is developed.

MDE Response:

- Maryland feels that this range is too broad. As suggested by the commentators, it would allow lakes currently classified as oligo-mesotrophic in the State's 305(b) report (see response to previous comment) to degrade to the brink of eutrophy.
- As refinements to the interim interpretation are incorporated, Maryland will consider the commentators' suggestion of altering this method of subdividing the mesotrophic range.

Comment 11. The document indicates that the minimum dissolved oxygen in the hypolimnion will be determined using published nomograms such as those of Reid (1961) or “comparable calculation methods”. We suggest that there be some criteria for determining what methods are “comparable” and acceptable.

MDE Response:

- Evaluations of any proposed alternative methods will be conducted as needed, on a case-by-case basis.
- MDE's intent is to allow the use of a mathematical method (e.g., equation or spreadsheet) that results in the same determination as the nomogram. Change “comparable” to “equivalent”.

Comment 12. This interpretation of dissolved oxygen standards will be used for listing impaired waterbodies and conducting TMDLs. We suggest that the procedure include an

indication of specific criteria for designating a lake as impaired. A single exceedance of the minimum hypolimnetic oxygen guideline should not constitute impairment. The procedure should specify how many excursions of the guideline are allowable, and over what time frame. For example, both U.S. EPA (1997) and Commonwealth of Virginia (2000) assessment methodologies include threshold exceedance rates to determine attainment of designated uses. Waterbodies are considered to partially meet the use if 11% to 25% of samples exceed the standard, and to fail to meet the use if more than 25% of samples exceed.

The proposed procedure fails to consider both the frequency and duration of D.O. excursions. For example, a single, short-term excursion of the minimum D.O. level is clearly of far less concern (and does not warrant listing) than repeated excursions, or an excursion lasting for weeks or months. The procedure should specify criteria that will be used for listing decisions that include frequency and duration, and prioritize high-frequency, long-duration excursions. Also, no data beyond a certain age should be used.

MDE Response:

- Interpretation of available data is necessarily dependent upon professional judgment. Water quality managers must consider the magnitude and frequency of the exceedance, as well the quality of the data and the comprehensiveness of the sampling regime. As above, this is one reason for keeping this decision methodology in guideline—as opposed to regulation—form.
- As refinements to the interim interpretation are incorporated, Maryland will consider the commentators' suggestion of including threshold exceedance rates to determine attainment of designated uses.
- Maryland recognizes and appreciates the importance of using recent data. EPA guidelines stipulate that data used for 305(b) listing be no older than five years. For TMDL development, Maryland's policy is to consider all available data, and use the best readily available data.

Biocriteria

Procedures for Implementation of COMAR § 26.08.02.03(C): Biological Assessment of Water Quality

Introduction

The Maryland Department of the Environment (MDE) solicited comments on the *Interim Framework for Application of Biocriteria Based on the MBSS Program to the Water Quality Inventory and List of Impaired Waters* from May 21 to July 31, 2001. Subsequent to that solicitation, MDE has revised the title in accord with future plans to promulgate a regulation on biological assessment, providing additional clarification on the implementation of its narrative water quality standards. Other revisions have been made in response to comments received and the input of the Biocriteria Advisory Committee. A list of commentators, their affiliation, and the date they submitted comments is provided below.

List of Commentators

Author	Affiliation	Date
Cy Jones	Washington Suburban Sanitary Commission	6/18/01
James Stuhltrager	Mid-Atlantic Environmental Law Center	7/31/01
Patricia Gleason	U.S. EPA Region III	6/15/01
Steve Stewart	Baltimore County DEPRM	6/21/01
Jules J. Loos	Mirant Mid-Atlantic	7/31/01
James F. Stine	Constellation Energy Group	7/31/01

Comments and Responses

Comment 1. Biocriteria and multi-metric biological indices are important tools in the scientific evaluation of surface waters in Maryland and throughout the US. With documents such as U.S. EPA's RBPs for Use in Wadeable Streams and Rivers (1999, 2nd Edition), good technical guidance exists on many scientific aspects of biocriteria. However, the proper application of biocriteria in determining whether a stream segment or watershed is impaired for regulatory purposes is more difficult. The State's goal is to properly identify waters that are impaired due to water quality, not to develop a procedure, which has the potential to include waters that are not impacted, or that do not achieve a specific numeric IBI value due to other reasons.

MDE Response: This comment raises two issues with respect to the Biocriteria decision methodology: (1) the procedure should not identify as impaired waters that are not impacted, and (2) the impairment should be judged only on the basis of water quality, not other factors such as habitat impairments.

The first issue implies that the State should develop a "perfect" method that results in no errors. No such test is possible. Further, any test that was designed so that it would never identify an

unimpaired water as impaired, would almost necessarily result in a test that would not identify some waters as impaired although they actually were. In statistics these types of errors are known as Type I and Type II. As the chance of a Type I error decreases, the chance of a Type II error increases. MDE believes that its methodology as stated balances these two types of error such that there is some chance that a water may be identified as impaired when it is not, although that chance is small. The chance of missing an impaired water is also kept at an acceptable risk.

EPA participated with the Advisory Committee that helped the Department develop this methodology. EPA notes in response to this comment:

It is our opinion that the MDE procedure is not overly protective. The metrics were scored using the 10th and 50th percentiles of the reference site distributions. The Versar analyses have indicated that the threshold of 3.0 for the total IBI scores translates to a range of the 12th % to the 17th% for the different fish IBI calibration reference site data sets (coastal plain, Piedmont, highlands, all) (Roth et al presentation to Committee). Given the fairly wide range of reference site F-IBI scores in the MBSS dataset, EPA guidance recommends that a larger margin of safety might be used to set a threshold for impairment. For example, the 25th % of the reference site scores has often been used as a threshold for determining impaired sites when the range of reference site scores is wide or there is a question regarding the true quality of the reference sites. In terms of classification efficiency, the analysis that Versar presented to the Committee indicated that the current threshold of 3 for the FIBI classifies reference sites properly 85% of the time in the Coastal Plain, 89% of the time in the Eastern Piedmont, and 84% of the time in the Highlands. Degraded sites were classified correctly 66% of the time in the Coastal Plain, 93% of the time in the Eastern Piedmont and 88% of the time in the Highlands. Clearly, the threshold of 3 is not resulting in a large number of original reference sites being misclassified as impaired. In addition, including an estimate of uncertainty helps to further decrease the chance that sites that are not impaired will be incorrectly classified as impaired.

The second issue is one that may be effectively dealt with by the new regulations that EPA is proposing that will allow an impaired water to be identified based on factors other than water quality and will not require a TMDL for those waters. It is anticipated at this time, that such flexibility will be extended to the 2002 list of impaired waters. On this issue it is also worth noting that sometimes the cause of an impairment cannot be identified at the time of listing. Listing the water provides an opportunity to direct resources to determine whether the stressor is water quality based or due to some other environmental condition.

Comment 2. The Department understands the legal, regulatory and technical difficulties of placing a stream segment on the State's 303(d) list, and should be careful to develop decision criteria tools with little potential for falsely categorizing a water as impaired. As noted above, the numeric IBI criteria are not adopted standards and cannot be used in the 303(d) listing.

MDE Response: MDE is not restricted to using adopted numeric standards but has flexibility to apply the narrative standards to situations for which statewide numeric standards have not been developed. However, a promulgation more directly referencing these biocriteria methodologies is planned for adoption prior to the final submission of the List.

EPA notes in response to this comment: "The IBI criteria are a direct measurement of support or non-support of the designated uses, and therefore can be used in listing decisions. The designated uses are part of water quality standards, as stated in all current 303(d) listing guidance. In July 1991, EPA transmitted final national policy on the integration of biological, chemical, and toxicological data in water quality assessments. According to this policy, referred to as "Independent Application," indication of impairment of water quality standards by any one of the three types of monitoring data (biological, chemical or toxicological) should be taken as evidence of impairment regardless of the findings of the other types of data."

Comment 3. Overall, the draft decision criterion lacks sufficient detail to allow readers and affected parties to fully understand exactly how this approach would be used at locations across the State. For example, it is imperative that the revised document explicitly cites the technical reference that will be used to calculate the F-IBI and B-IBI indices, as there are numerous varieties of the IBIs and other multimetric indices that are used across the country. More specific comments are presented below.

MDE Response: Agreed. The final draft of the document will include a reference section citing the original peer reviewed documents upon which the decision methodology is based.

Comment 4. The procedure for evaluating 12-digit subwatersheds is flawed for many reasons. First, the procedure allows a stream segment to be listed as impaired based on a single sample (F-IBI or B-IBI) that has a low value.

MDE Response: The Advisory Committee had significant concerns about this issue and addressed it in detail. An analysis of replicate samples (samples of adjacent areas in the case of the F-IBI) found that in fact single samples were representative and that in most cases additional samples gave the same results. The confidence interval approach further takes into account the uncertainty embodied in a single sample. If the magnitude of deviation from an acceptable value is sufficiently large, and there is no apparent reason (such as mishandling of a sample) other than water quality why it should be so far below an acceptable value, it is reasonable to conclude that there is a problem even on the basis of a single sample. The proposed methodology further quantifies that uncertainty at the 10% level (i.e., we are 90% confident that the sample value is lower than the acceptable value of 3).

Comment 5. There is no allowance for additional samples to be collected and evaluated before the segment is listed as impaired, which would inflict hardship on any facility located on that subwatershed.

MDE Response: There is every allowance to allow additional samples to be collected before listing, resources simply do not allow that to happen. More data is always preferable, but EPA's directions are explicit on this: available data are to be used as the basis for the decision. In most

cases, significant data will be collected prior to TMDL development, which data may provide a basis for delisting the waterbody.

Comment 6. There should be an explicit statement that more data are better than less data, and that a watershed determined to be impaired based on a small number of samples can be easily removed from the State's 303(d) list based on newer and better data.

MDE Response: See response above. As a general rule, new data from the same location as the data used to list a waterbody showing that the waterbody is no longer impaired would generally be adequate for delisting.

Comment 7. Further, the Department should be required to notify all dischargers to a waterbody which is to be listed as biocriteria-impaired, detailing the implications of the listing, and giving them sufficient time to conduct additional studies during appropriate seasons. We recommend that at a minimum 2 out of 8 samples indicate impairment, and that a procedure for closely examining the failed samples be developed to determine if these failures indicate true impairment. Also, no data beyond a certain age should be used.

MDE Response: There does not seem to be any basis for treating a biocriteria-impaired waterbody differently from a waterbody impaired for any other reason. This suggestion that 2 of 8 samples be used to indicate an impairment is not possible to achieve when multiplied across every waterbody in the state. Further, a frequency of impaired samples does not take into account the magnitude of the exceedance nor the variability around these samples. The proposed methodology accounts for both. In general, DNR in the 305(b) report and MDE in the 303(d) list follow that guideline, but it is not a rule, and depending on data availability and method used to derive the data, older data may be used if it is still believed to be representative.

Comment 8. The "rules" for statistically determining whether a value results in meeting or not meeting criteria needs to be written more clearly. For example, does "*the value and the upper bound*" mean that both must be greater than or equal to 3? Further, if the [single] value is greater than 3, then the upper confidence limit must be greater than 3.

MDE Response: We will revise the document and provide figures to clarify these points.

Comment 9. The Department's proposed method allows a criteria compliance determination based upon a single F-IBI or B-IBI score (see comment above). But because statistics cannot be conducted on a single value, the Department's approach requires that confidence limits be calculated around that single value based upon a fixed coefficient of variation (CV=0.08, which seems implausibly low) determined statewide by MDNR.

MDE Response: EPA responds: This value is not implausibly low, and it was developed using an acceptable statistical method that is similarly outlined in the national RBP guidance, EMAP methods, and elsewhere. This estimate also is in good agreement with other estimates we have developed on our own data and seen in other data sets. In the revised RBP guidance the coefficient of variation for an index in Wyoming was estimated at 2.9%. For a project in West Virginia we estimated the standard deviation of repeated measures for an index on a 1 to 100

point scale as 4.3. The overall CV for the dataset was estimated as 7%. This is similar to the standard deviation for repeated measures and so that the state of West Virginia estimated with an independent statewide data set, using the same multimetric index.

Comment 10. As discussed above, a waterbody should never be listed as impaired based on a single value, or the ability to generate additional data to adequately demonstrate that the segment is impaired.

MDE Response: The current 303(d) listing guidelines do not agree with this comment. Further, MDE has determined that the proposed approach does adequately demonstrate that the segment is impaired.

Comment 11. The Department's approach requires that confidence limits be calculated around that single value based upon a fixed coefficient of variation (CV=0.08) that was determined statewide by MDNR. Using this approach, however, assumes that the variability around a single fixed value is *constant* regardless of the experience of the researcher, the type of sample, the season of the year, the habitat and location of the sample, etc.

MDE Response: This estimate was derived for the MBSS data set, which was collected by trained field teams, using the same methods, sampling the same habitat, in the same index period. There is no assumption that the CV is constant, rather the default CV is an average that is used when site-specific data are not available, and is a reasonable estimate of the true population CV.

Comment 12. If it is used in the Department's assessment approach, the Department needs to carefully support the proposed CV=0.08 value and its *universal* applicability to Maryland waters.

MDE Response: The CV does not have a wide range, even when it is broken out by condition categories or stream order and is a rational approach to estimating variability in the absence of replicate samples.

Comment 13. The last paragraph of Section 2 states that these site data are "only a snapshot of conditions" and "does not necessarily reflect chronic conditions". We disagree in that one of the long-recognized strengths of benthic macroinvertebrate evaluations is that the community reflects longer-term conditions than chemical or toxicological sampling (often reflecting 6 months to more than a year).

MDE Response: We agree, and the scientific community agrees that biological data are representative of longer time frames. We will revise the document to reflect this.

Comment 14. The term "habitat" is nearly absent from the Department's draft, yet its importance in the understanding of specific biocriteria values cannot be argued. U.S. EPA's recent RBP manual (Barbour et al 1999) explicitly states:

The procedure for assessing physical habitat quality presented in this [EPA] document is an integral component of the final evaluation of impairment (p. 3-4). Habitat, as

structured by instream and surrounding topographical features, is a major determinant of aquatic community potential (p. 3-5). Where physical habitat quality at a test site is similar to that of a reference, detected impacts can be attributed to water quality factors (i.e., chemical contamination) or other stressors. However where habitat quality differs substantially from the reference conditions, the question of appropriate aquatic life use designation and physical habitat alteration/restoration must be addressed. Final conclusions regarding the presence or absence of biological impairment should thus include an evaluation of habitat quality to determine the extent that habitat may be a limiting factor (p. 3-5).

Neither the B-IBI nor the F-IBI calculations proposed for use by the Department include any evaluation of the quality of the habitat from which the samples were collected.

MDE Response: Evaluation of the habitat data will help diagnose the causes of the biological impairment and is part of that diagnostic step in the analysis of the data. This is part of the stressor identification that is referred to, but not addressed in detail by this decision methodology.

Comment 15. For this reason, an IBI score could be less than 3.0 (even though water quality is excellent) solely because the natural habitat of the sampled area is poor.

MDE Response: This goes to stressor identification, not to a decision on whether the waterbody is impaired. With the new flexibility in listing using a five-part list it will be easier to address this issue.

Comment 16. Further, because the IBI value of 3.0 is based upon statewide and regional "reference" areas, few urban stream segments will ever comply with the Department's proposed biocriteria; many will be formally listed as impaired, and may require unknown changes to be made in permits throughout the watershed. Habitat is clearly an important factor. No stream segment should ever be listed on the State's 303(d) list without a formal habitat evaluation to determine whether the stream segment is meeting its potential biological condition. This is particularly important for urban streams that have permanently changed conditions that cannot be retrieved.

MDE Response: This is an issue the Committee has already discussed - the role of Use Attainability Assessments in the regulatory process. Urban streams would have to be redesignated using a UAA to have a reference condition specific to urban streams. The urban streams would have to be redesignated to a use that may not meet the interim goals of the CWA.

Comment 17. The approach should specifically address the selection of appropriate reference streams for urban and developed drainage basin conditions. Special indices should also be developed for urban streams, as is proposed for other special condition areas such as limestone streams.

MDE Response: Again, this could only be done if a UAA was completed and approved. As we have heard in the Committee, some communities may agree with redesignating urban streams,

some may not. This would have to be a process involving the entire public. There is a difference between the case of urban streams and limestone streams. One can argue that limestone streams are unique naturally and therefore deserve their own reference condition to support a general aquatic life use in these streams. For example, blackwater and unproductive cold water streams are also singled out in the exceptions as unique naturally. Urban streams are not naturally unique. They have been changed (in some cases irreparably) by human activities. Therefore, it is more appropriate to use the UAA as a way to address irretrievable impairment, or incorrect original designated uses (or the existing use).

Comment 18. Section 3C states that if a watershed is determined to be impaired, then "the cause of the impairment(s) will then be determined." The text should be rewritten to say that the potential cause(s) of the impairment(s) will be *investigated*. Accurately determining the cause(s) of a F-IBI value being 2.6 instead of 3.0 will be very difficult in most instances. Further, it will be even more difficult Department staff to determine what specific improvements will be required for each point- and non-point source discharger to a "biocriteria-impaired" segment to achieve an IBI score of 3.0. Has the Department determined how this can reasonably be done in waters throughout the State? For example, if the impairment (as reflected in low IBI scores) is primarily caused by poor habitat, how will dischargers be affected by that decision, and what future permit constraints will they face? What strategy will the Department use to improve a specific IBI score for a watershed if the cause(s) of impairment cannot be determined?

MDE Response: All good comments. The ability to model biological response is clearly not developed to the point where the predictions can be made. However, it is part of the TMDL process to collect more data, and to predict what needs to be done to meet water quality standards.

Comment 19. The last sentence in the last paragraph of Section 4 does not make sense as written.

MDE Response: We will review and revise as appropriate.

Comment 20. There is no indication that the Department has a workable plan for how biocriteria-listed stream segments will be addressed in terms of stream restoration (or Individual Control Strategies) for facilities in an affected watershed.

MDE Response: There are several State and local funds including Section 319 funds, Small Creeks and Estuaries funds, Chesapeake Bay Implementation Grants, local capital projects etc. that can be used for stream restoration. The presence of these "biocriteria-impaired" streams on the 303(d) list can provide a foundation for priority setting for restoration.

Reference used by commentator:

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribbling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates

and Fish. 2nd Edition. EPA 841-B-99-002. U.S. EPA Office of Water, Washington D.C. [<http://www.epa.gov/owow/monitoring/techmon.html>].

References used for responses:

- Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribbling. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish. 2nd Edition. EPA 841-B-99-002. U.S. EPA Office of Water, Washington D.C. [<http://www.epa.gov/owow/monitoring/techmon.html>].
- Roth, N. et al., Answers to IBI Questions (Presentation to Biocriteria Advisory Committee). Versar Inc. Columbia MD.
- Roth, N. et al. Biological Indicator Variability and Stream Monitoring Program Integration: A Maryland Case Study. Versar Inc. Columbia MD. October 2000. (In draft)
- USEPA. EPA's Policy on the Use of Biological Assessments and Criteria in the Water Quality Program, Washington, DC. May 1991.
- USEPA. Guidelines for Preparation of the Comprehensive State Water Quality Assessments (305(b) Reports) and Electronic Updates: Report Contents. EPA-841-B-97-002A. Washington DC. September 1997.
- USEPA Region III. Guidance for the Prioritization and Targeting of Waters Listed Under Section 303(d) of the Clean Water Act and Determination of Existing and Readily Available Water Quality-Related Data and Information for Listing Waters Under Section 303(d) of the Clean Water Act. Philadelphia, PA. July 1997.

Comment 21. Maryland proposes that a water will be identified as impaired if it (1) does not meet either the F-IBI or B-IBI criteria and (2) a “review of other biological data” demonstrates impairment. This two-part test is flawed because it does not identify what other biological data will be reviewed and because the IBI criteria provide a reliable methodology without the need for additional review.

MDE Response: This caveat is inserted simply to provide a “reality check.” If other biological data are readily available and strongly contradict the IBI results, State biologists would seek to identify the cause of the contradiction and determine which results are more reasonable. If no data are available, or if other biological data are consistent with the IBI findings, the waterbody would be listed appropriately. This is not a specific requirement to find more data, it meets the requirement to evaluate all available data.

Comment 22. The Interim Framework provides a reasoned approach to the listing of waters based on biocriteria. The inclusion of the cause/source identification (3.C.) procedure is important in order to correct any impairments to the biological community.

MDE Response: No response.

Comment 23. Where interested parties have undertaken biocriteria assessments (for example, in estuarine waters), MDE should have the flexibility to consider such information in the listing process. We recommend amending the framework to provide this flexibility.

MDE Response: The listing framework specifically refers to the application of the MBSS data which applies to streams. This does not in any way preclude the Department from considering other data. However, at this time the Department will consider such data on a case-by-case basis. The Biocriteria Advisory has discussed the need to develop additional frameworks, such as for estuarine areas.

Comment 24. It is inappropriate to list waters impaired due to natural conditions or other causes unrelated to the discharge of pollutants. In some cases alteration in the biological community can be traced to causes other than pollutants such as habitat destruction. Therefore it is imperative that the cause of the impairment be determined before a waterbody is listed.

MDE Response: It is often difficult to demonstrate and justify a natural conditions exception and that information frequently cannot be obtained in the timeframe and with the resources available for the 303(d) listing. If such information were to be available it could certainly be considered. The new structure of the 303(d) list provides more flexibility in response to this type of situation, and will be used appropriately.

Bacteria

Introduction

The Maryland Department of the Environment solicited comments on the Decision Rules for Listing Waters on Maryland's 303(d) List from May 2 to July 31, 2001. A list of commentators, their affiliation, and the date they submitted comments is provided below.

List of Commentators

Author	Affiliation	Date
C. Victoria Woodward	Safe Waterways in Maryland	5/30/01
Patricia Gleason	U.S. EPA Region III	6/15/01
Steve Stewart	Baltimore County DEPRM	6/21/01
James M. Stuhltrager	Mid-Atlantic Environmental Law Center	7/31/01
Will Bullard	US Navy/DOD Regional Environmental Coordination	8/9/01
Cy Jones	Washington Suburban Sanitary Commission	6/18/01
Meosotis Curtis	Montgomery County	6/14/01
Guy Tilghman Hollyday	Baltimore Sanitary Sewer Oversight Coalition	6/13/01
Paul Sturm	Center for Watershed Protection	6/21/01
Loretta Richardson, Pat Lane	League of Women Voters of Baltimore	6/28/01

Comments and Responses

Comment 1. Under “Interpretation of fecal coliform data in Use II, Shellfish Harvesting Areas” all areas where known and persistent failures occur should be listed. Cannot avoid listing with a buffer area.

MDE Response: The actual monitoring data provide a more accurate assessment of conditions than the number of spills because shellfish areas are monitored twice per month and meet the standard. It is very unlikely that areas subject to persistent failures would be able to meet the very stringent shellfish requirements.

MDE does not assume that a buffer area is sufficient to avoid a listing. Restricted areas around WWTP outfalls near shellfish beds are additional protection to the shellfish consuming public; areas must still meet bacteriological standards.

Note that this question applies exclusively to Use II waters and not generally to other waters.

Comment 2. Under “Interpretation of Fecal Coliform Data for Use I, III, or IV Waters” a technological fix does not preclude a listing decision.

MDE Response: MDE will add the following change:

Interpretation of Fecal coliform Data for Use I, III, or IV Waters

a) Those segments impacted by faulty sewer lines, excess inflow and infiltration, or other impacts requiring a technological fix will be listed only if the problem cannot be fixed before the next listing period.

b) Those segments where no technological fix is feasible, or where that repair cannot be completed by the next listing period, or where there is a potential human health risk, are listed.

Comment 3. This document needs to include the bacteriological standards of E. coli and enterococci.

MDE Response: The new indicators currently apply at permitted beaches, but have not yet been promulgated for other designated uses, thus there is little or no E. coli and enterococci data to evaluate for 303(d) at this time.

Comment 4. Sanitary surveys for source identification have been given greater weight than the numerical standard. It should be clarified who will be conducting the sanitary surveys.

MDE Response: We do not believe that greater weight has been given to sanitary surveys than to the numerical standards, both are important. Sanitary surveys are conducted by both MDE and County Health Departments in response to the “Needs Survey” which finds failing septic systems, prioritizes areas for corrections of those failures and/or connection to sewage lines.

Comment 5. “Technological fixes” is ambiguous. Example: if there is a bacteriological impact due to wash-off from storm events, does a stormwater BMP constitute a technological fix?

MDE Response: If an impact in ambient water due to bacteriological water quality is identified and can be corrected by fixing or retrofitting a stormwater BMP quickly enough to demonstrate before the next listing period that there is no longer a bacteriological impact that would apply. The key point is fixing the impact, without the need for a TMDL because the source of the impact is clearly identified and can be remedied quickly.

Comment 6. Maryland must list waters that are closed to harvesting following rainfall greater than one inch in 24 hours. §303 does not allow such an exception.

MDE Response: These areas are not closed because there is an actual water quality impairment, but for public health reasons because there is a chance that there could be brief exceedances. Shellfish can concentrate bacteria or pathogens that may be present in very low numbers, and because they are eaten raw, exceptional measures must be taken. It is likely that in most of these cases, monitoring shortly after the rainfall would show no impairment. That is irrelevant to the public health concern, but very relevant to the 303(d) listing.

Comment 7. MDE incorrectly interprets the water quality standard by saying that an exceedance of the standards does “not necessarily represent an impairment.” A public health hazard is presumed if the numeric criteria are exceeded.

MDE Response: The commentator neglected part (c) of that paragraph which notes that numeric criteria do not apply “when a sanitary survey approved by the Department discloses no significant health hazard.”

Comment 8. Beaches with “chronic closures” are listed. Chronic is ambiguous, what is the basis for this determination?

MDE Response: A 10% closure period is the break point separating chronic (long-term) and acute (short-term) swimming area closures. Swimming area closures may be a single, long-term event or multiple short-term events. A 90-day swimming season (Memorial Day to Labor Day) is assumed. A beach with closure(s) exceeding 10% of that time - a cumulative total of 9 days or more during a year's swimming season - is listed as chronic. The closure counter is reset each calendar year. All closure events reported to the US EPA (documented in their annual "Beach Watch" report) or to the Natural Resources Defense Council (documented in their annual "Testing the Waters" publication) are listed in the State's 305(b) report. Closure events out-of-season are reported. Most closure incidents are a result of a spill or other short-term pollution event that is monitored and addressed.

This process has been used for more than a decade and developed prior to EPA's guidance, which suggest using a one week period as the separation between partial and not supporting classifications. EPA's guidance covers the entire country where the swimming season varies from 0 days to year-round. Their standardized 7-day period does not reflect local conditions or issues that often result in reported beach closures.

The rationale for this decision-making process is that a swimming area closure represents an agency response to possible human health risk - closures are not a direct measure of water

quality conditions. Such data usually is not available. Some beach area closures have no water quality data associated with them - closures may be made for any number of reasons by local health officials. Closures reported for reasons other than poor water quality condition (e.g. sanitary concerns - drinking water reservoirs; physical safety concerns - lack of lifeguards, marinas, Inner Harbor, areas of swift currents or hidden rocks/holes) are not reported if they can be identified.

Closure events over a significant part of the use season (10 percent) identifies long-term water quality issues that may be reflected socially/economically in terms of the intention/behavior of bathing uses by the local population as well as those traveling to visit an area.

Comment 9. Do approved shellfish waters always meet water quality standards or is there some allowance for infrequent exceedance?

MDE Response: The National Shellfish Sanitation Program criteria, which apply in shellfish waters look at a long-term (at least 30 values) measures such as the median and 90th percentile, not at individual data values. Both of these measures allow for infrequent exceedance.

Comment 10. Does “Bay Water Quality Tributary Monitoring” imply that only Bay Program monitoring data will be used?

MDE Response: Bacteriological monitoring is done by the field teams that also do the Bay Tributary Monitoring and the implication was a misstatement. Routine monitoring is conducted in major tributaries, and all relevant data are used.

Comment 11. Will a single exceedance of the geometric mean criterion trigger a sanitary survey?

MDE Response: Sanitary surveys are usually done on a routine basis to identify and fix broken sewage infrastructure or failing on-site disposal systems. If a sanitary survey had been conducted in the area where there was a single exceedance of the geometric mean, and if the sanitary survey revealed no problems, MDE would not consider the single value as evidence of impairment. If the geometric mean was routinely exceeded, a sanitary survey would be conducted, resources permitting, to see if a technological fix was possible, otherwise an impairment listing is likely.

Comment 12. Compliance issues and listing of priority waters, presumption of permit compliance.

MDE Response: Generally, MDE finds few violations of the fecal coliform limits in treated effluent. If there is evidence of violation of fecal coliform permit limits, compliance and enforcement will be used to correct the problem quickly. If for some reason, the violation cannot be quickly corrected, segments will be listed. Aging collection systems and, in a few towns, combined sewers may contribute fecal coliform to surface waters. Non-point sources such as run off from land uses ranging from woodland, marshland, rural, suburban, urban, animal agriculture, etc. are considered, and such waters are listed when discovered. Treated effluent

from WWTPs or other permitted dischargers is generally not a major contributor to fecal coliform in surface waters.

pH

List of Commentators

Author	Affiliation	Date
Patricia Gleason	U.S. EPA Region III	8/8/01

Comments and Responses

Comment 1. EPA Region III has reviewed the draft methodologies for sediment and pH for listing pollution impaired waterbodies on the 2002, 303(d) list, as published on MDE's web site July 2, 2001. Based on that review, we have the following comments.

Regarding the pH methodology, the exclusion of streams from the 303(d) list due to atmospheric deposition events is not correct. EPA Region III's 1997 Guidance for the 303(d) list (Enclosure–page L12), indicates that streams should be listed as impaired if the impairment is due to atmospheric deposition. Section 40 CFR 130.7(b)(1) does not differentiate between waters being impaired by land based point and/or nonpoint pollutant sources or atmospheric deposition. Although controlling pollution from atmospheric deposition can be very difficult, Section 303(d) of the Clean Water Act (CWA) and implementing regulations at 40 CFR 1307.7 do not require that the decision to include a water on a state section 303(d) list depend upon the ease with which a source of a pollutant can be controlled.

Therefore, states should include on their section 303(d) list, waters that are impaired by pollutants entirely or partially from atmospheric deposition, regardless of the source of the pollutant or the ease with which the pollutant can be controlled. States may wish to assign a lower priority to such waters. In the near term, States may want to consider the extent to which existing air pollution control authorities such as State implementation plans and local ordinances, could be used to address the listing of waters impaired due to atmospheric deposition. However, in those cases where atmospheric deposition is associated with long-term transport and sources and effects are not well understood, it is appropriate to view listing as a starting point for the State and EPA to collaborate on a longer-term strategy for scientific research and remediation.

MDE Response: The following line will be added to the text.

Waterbodies displaying acidic conditions as a result of atmospheric deposition will be placed on the 303(d) list if it is determined that there is not adequate natural buffering capacity in the watershed.