January 19, 2011

Subject:
Comments on the Scientific Advisory Board 12-28-2010 Review of EPA Draft Report, March 2010, entitled "A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams"

The National Mining Association (NMA) is pleased to provide these comments on the EPA's Science Advisory Board (SAB) draft report dated Dec. 28, 2010 that provides a review of the EPA draft report entitled; *A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams* (EPA 2010). The SAB generally supported the methodology and the conclusions, but identified significant recommendations for improvement of the scientific validity of the results. The SAB report responded to 8 questions posed by EPA concerning the Conductivity Benchmark Report. NMA would like to address a few of NMA's primary issues and concerns with the SAB report.

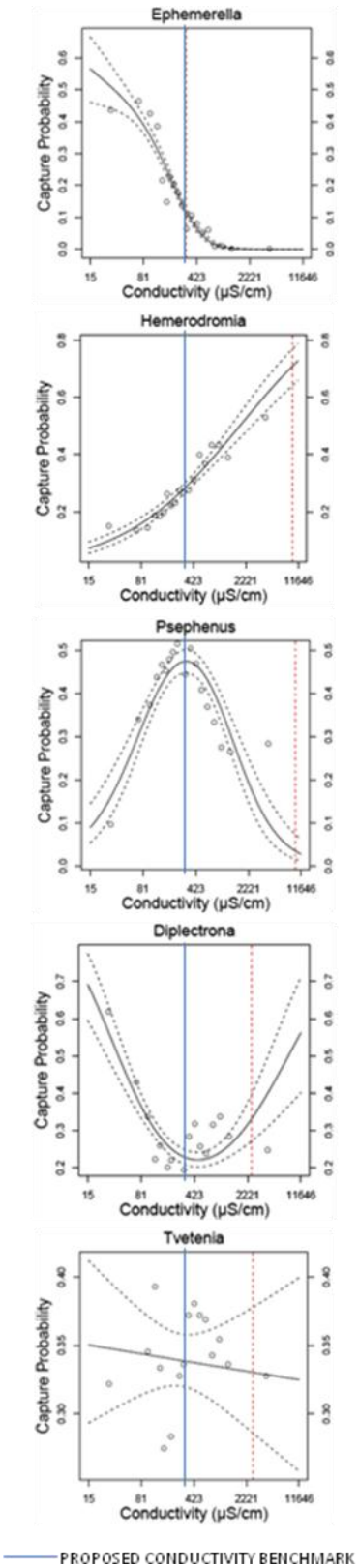
Charge Question 2: The derivation of a benchmark value for conductivity was adapted from EPA's methods for deriving water quality criteria. The water quality criteria methodology relies on a lab-based procedure, whereas this report uses a field-based approach. Has the report adapted the water quality criteria methodology to derive a water quality advisory for conductivity using field data in a way that is clear, transparent and reasonable?

The SAB report concluded that the use of a field-based approach to developing the benchmark was justified, but based their comment from the perspective that it was better than water quality criteria derived in the laboratory using *Ceriodaphnia* or similar traditional laboratory toxicity test species. The report states that the approach provides an improvement over a benchmark that might have been derived from laboratory data using test species that are not native to the region and do not reflect the broad range of life stages.

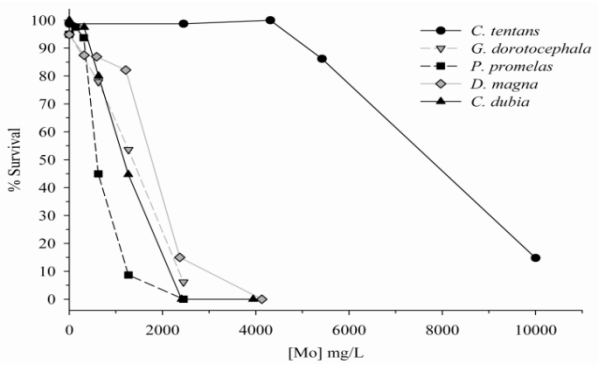
However, both the EPA report and SAB note that the laboratory testing approach has been successful and most amenable to toxicants (e.g., ammonia, metals) with clear and consistent modes of effect. The traditional method for establishing water quality criteria relies on information from controlled laboratory experiments to identify a direct relationship between organisms' response and exposure concentrations of a water quality constituent without the confounding influences.

Species Sensitivity Distributions (SSDs) are supposed to be generated from laboratory data based on actual causal relationships and not from field biomonitoring studies based on correlations. The SAB is inconsistent in their criticisms of the use of laboratory testing of sensitive species such as mayflies for potential toxicity associated with elevated conductivity, and their recommendation to include citations of laboratory studies in the literature showing the physiological mechanism and effect between concentrations of dissolved ions and other stress factors.

The SAB report identified numerous other caveats on use of a field based water quality methodology. The SAB recommended that the stressor-response relationships be evaluated using multiple analytical approaches. They felt that the methodology could only be transferred if there was a large, robust data set, which included both stressed and reference sites. As confounding factors increase in a study, the data set size would also need to increase. In addition, they noted that the benchmark should be backed up by applicable lab tests, reverting to the value of the laboratory-based study. The use of the field methodology should examine mode of effect, and address confounding factors in detail. The SAB also desired to see an analysis of the impacts on other components of the aquatic community such as fish, amphibians, or long-lived macro-invertebrates such as mollusks.



In addition, the SAB has continued to ignore significant issues raised by GEI Consultants, in a report for NMA, specific to how EPA used conflicting stressor response profiles in their SSD derivation. As outlined in GEI (2010), the prime underlying principle governing the use of a SSD is that all of the organisms in the distribution exhibit a consistent response to the stressor. Specifically, each of the taxa should respond negatively to the stressor – only differing in their degree of sensitivity – as shown below (Canton et al. 2010).



However, as illustrated to the left (from EPA 2010), there are five different ways the organisms used to derive EPA's benchmark would seem to respond to conductivity. These differences do not reflect varying levels of sensitivity over a consistent response profile – *rather, they are fundamentally different types of stressor-response profiles*.

These five stressor-response profiles provide substantially different answers to the question "*what conductivity concentration is necessary to provide the level of protection used by EPA?*

- Decreasing (*Ephemerella*): <300 µs/cm
- Increasing (*Hemerodromia*): >300 µs/cm
- Optimum (*Psephenus*): >75 <u>and</u> <2,500 µs/cm
- Bimodal (*Diplectrona*): <200 <u>and</u> >2,000 µs/cm
- No response/bimodal (*Tvetenia*): *None needed*

There is no way to reconcile these widely conflicting stressor-responses into a single benchmark protective of the entire macroinvertebrate community – something not taken into account by the SAB. Indeed it is even more probable that capture probability of species other than those that respond in a decreasing fashion to conductivity are a result of ecological factors other than conductivity, such as habitat, flow, food resources, etc. SAB continues to downplay the significance of these factors.

SAB noted several areas where the clarity and justification of the approach and benchmark could be enhanced. The SAB felt that the study could be improved by justifying the use of conductivity as an indicator, rather than particular ions or ion ratios, and relating conductivity to other aquatic effect endpoints (species other than mayflies). While the EPA study found a moderate correlation between conductivity and a number of *Ephemeroptera* genera there is still a wide scatter in the relationship, due to confounding factors which made interpretation of the data difficult. This together with the relatively high conductivity LOEC (lowest observable effects concentrations) values in laboratory studies casts considerable doubt on conductivity (or the ions measured by conductivity) as the dominant causal factor for the number of mayfly genera observed at sites included in the study.

The report has not developed a water quality criteria field-based methodology to derive a water quality advisory for conductivity that is clear, transparent and reasonable, and that follows long-standing EPA criteria derivation protocols.

Charge Question 3: Appendix A of the EPA report describes the process used to establish a causal relationship between the extirpation of invertebrate genera and levels of conductivity. Has the report effectively made the case for a causal relationship between species extirpation and high levels of conductivity due to surface coal mining activities?

The EPA report has not made a strong case for the relationship between species "extirpation[1]" and high levels of conductivity due to surface coal mining. While EPA linked stream conductivity and the amount of valley filling of the watershed, and associated it with effects on benthic macroinvertebrate communities generally, the

[1] Note the use of the term "extirpation" is incorrect, based on how EPA presents data. In fact, the endpoints are simply related to changes in "capture probability". Extirpation is a process that results from a series of biological factors that lead to the local loss of a taxon from a particular stream or region. To our knowledge, such a process has not yet been demonstrated as a result of elevated ions related to conductivity.

SAB noted specifically that conductivity is not a "pollutant". Elevated conductivity or dissolved solids concentrations are known to be associated with disturbed landscapes, not exclusively mining land uses. The Causal Assessment in Appendix A was based on reducing 45 land use types down to 7, of which only one is Barren Land (Appendix F). The original land use types included other disturbed land uses, including unpaved roads, oil and gas related disturbance, logging (harvested forest and skid roads), burned forest and recent urban development.

Many of the potential causes for biological impairment can act together to cause impairment. EPA's Causal Analysis/Diagnosis Decision Information System (CADDIS) also provides recommendations and cautions in the application of statistical tools. It states that statistical significance does not equate to biological significance. Statistical analysis does not tell us whether variability in observations is caused by the stressor being analyzed or whether variability in observations is biologically relevant.

CADDIS cautions that "*Concluding that a candidate stressor is or is not the cause based on a correlation coefficient is inappropriate because:*

- *Stressors often covary with each other and with natural environmental attributes. A strong relationship between the biological response and candidate cause could reflect a covarying stressor or natural factor other than the candidate cause,*
- *Hypothesis testing was designed for interpreting controlled experiments with replicates and random assignment of treatments, and*
- *Field data from observational studies rarely include replicates and "treatments" are not randomly assigned, therefore even strong associations do not prove causation.*

These cautions in CADDIS have not been adequately addressed in the EPA Conductivity benchmark study nor by the SAB in its review. It is difficult to understand why EPA failed to utilize their established scientific approaches to this conductivity benchmark study. Experimental validation, using mesocosms or other whole ecosystem experimental approaches, have been widely used in the past to confirm causality. We see no reason why the SAB should not also insist upon such experimental validation, and instead support the weaker and less direct statistical approach for establishing causality that EPA used.

The SAB appears to rely on the literature that for some aquatic organisms, stress can occur when ion concentrations in water exceed certain thresholds. However,

the SAB fails to compare these individual ion levels with the conductivity benchmark. The issue is not whether ions at some levels can cause stress but rather at what conductivity values, ionic compostion, and exposure time periods does conductivity result in chronic toxicity to sensitive aquatic organisms, and which organisms are truly the most sensitive. The SAB repeatedly makes the claim that *Ceriodaphnia* are not the most sensitive species with respect to ion stress despite the few and contradicting studies with mayflies which they conclude are more sensitive. Given the substantial literature demonstrating the sensitivity of *Ceriodaphnia* to ion toxicity, this claim is as yet poorly supported by the SAB.

Charge Question 4: In using field data, other variables and factors have to be accounted for in determining causal relationships. Appendix B of the EPA report describes the 11 techniques for dealing with confounding factors. Does the report effectively consider other factors that may confound the relationship between conductivity and extirpation of invertebrates? If not, how can the analysis be improved?

EPA's analysis of confounding factors is insufficient and does not represent a rigorous analysis of the true factors affecting the number of Ephemeroptera genera (and certainly not other insects or invertebrate groups) at sites included in the study. Other conditions including channel alteration, the presence of upstream impoundments, dissolved oxygen, and temperature represent well-documented significant factors that need to be evaluated. The causal analysis of biological impairment needs to be performed at each site because multiple stressors can act together to cause impairment and these causal factors vary considerably among sites.

The SAB review elaborated on the limitations of traditional laboratory methods but failed to address the primary limitation of the field-based approach followed in the benchmark study. That limitation is that it is based on a Species Sensitivity Distributions (SSD) derived from the data set based on the assumption that conductivity is the primary and only factor responsible for the local "extirpation" (really, reduced capture probability) of aquatic genera. In its response to **Charge 8** concerning the potential application of the field-based approach to other pollutants, the SAB states that the SSD methodology was readily applicable to conductivity because there is a relatively direct physiological mechanism and effect between conductivity, as a surrogate for concentrations of dissolved ions, and the occurrence of taxa. However, other stressors, such as nutrients and physical

habitat measures, resulted in effects which are moderated by other (confounding) factors. The confounding factors mentioned by the SAB include shading, habitat, base flow, bedded sediments and channel modifications. The SAB does not address the inconsistency in its caveats concerning these confounding factors for other pollutants but not for conductivity.

The SAB also did not acknowledge NMA's consultant, GEI's, independent statistical analysis of the dataset[2] which indicated that substrate composition, ionic composition, and channel features were better predictors of benthic macroinvertebrate community structure than use of conductivity alone.

Charge Question 5: Uncertainty values were analyzed using a boot-strapped statistical approach. Does the SAB agree with the approach used to evaluate uncertainty in the benchmark value? If not, how can the uncertainty analysis be improved?

Uncertainty analysis using a boot-strapped statistical approach does not adequately address uncertainty in the approach and the benchmark value. The SAB review commends EPA for initiating an uncertainty analysis, the authors note that there are other uncertainties in the benchmark that are not assessed using the bootstrap resampling procedure. One example is the uncertainties in the assignment of cause and effect between specific conductance and macroinvertebrate "extirpation" (i.e., reduced capture probability).

The largest source of uncertainty is in the application of the methodology and assumptions, particularly in the assumption that conductivity is the variable controlling the "extirpation" (i.e., reduced capture probability) of aquatic species in the study area. Figure 7 of the EPA document shows a considerable range in the XC_{95} values that occur just due to the variation in sampling from the same West Virginia data set in the bootstrap analysis. For example, the XC_{95} value calculated for the caddis fly genus *Lepitostoma* varies widely from less than 100 to over 700 µS/cm, due to the random variation in the sampling points as reflected in the bootstrap analysis. This high variability in XC_{95} values calculated for each genus in the data set clearly demonstrates that conductivity is not controlling the "extirpation" of aquatic species for most of the sampling points in the study area.

[2] It should be noted that the draft version of the GEI report that was submitted in advance of the SAB meeting in July 2010 did not contain a complete version of this statistical analysis. The complete version of this analysis is presented in GEI's final report from September 2010.

The conductivity benchmark value is derived from species SSDs developed from field data based on the poorly supported assumption that conductivity is the only (or at least the primary) factor responsible for the capture probability of sensitive aquatic genera that is represented by the field based SSD. If this approach is valid, why do the SSD plots for two sensitive genera (*Leptophlebia* and *Ephemerella*) presented in Figure 3 of the SAB report show such variability between the benchmark study and the results from Ohio? The SAB suggests that XC_{95} values for these genera may be higher in Ohio but fails to recognize that these results show the limitations of the approach for determining a conductivity benchmark or chronic toxicity value from field data. If conductivity were indeed causally related to the capture probability of individual taxa in the field, their XC_{95} values should not appreciably depend on what state the populations are located.

The uncertainty analysis does not address many of the other assumptions that were made. For instance, a genus was excluded from the "extirpation" calculation if it was observed at less than 30 of the 2,145 sampling sites. Perhaps a more appropriate threshold for exclusion is if it was not observed at more than 10% of the sites. This would eliminate the genera that are not commonly observed in the study area.

In conclusion, EPA aquatic life criteria are designed to protect the overall integrity of an aquatic community, rather than all of the species at all times. It is imperative to evaluate confounding factors and specific ionic concentrations to establish the specific cause, rather than using a water quality criteria surrogate of conductivity. Field testing should be backed up by laboratory studies documenting test duration, species, ionic composition and hardness. The use of conductivity as a surrogate parameter detracts from identifying the source of the problem. This will result in unanticipated consequences such as false positives for impairment. Industry and the state agencies will need to meet the conductivity benchmark and conduct TMDLs when the benchmark is not attained even though many unimpaired streams have conductivity levels above the benchmark. The increased regulatory burden resulting from false positives would financially burden the states; businesses would limit investment, which in turn would decrease local economic development. Specific ions most directly related to toxicity are better suited for regulatory criteria, facilitating industry's establishment of environmental goals, and government's interpretation of compliance.

Thank you for your time and attention to the National Mining Association comments on this important issue. Please feel free to contact me with any questions or concerns at (202) 463-3240 or at kbennett@nma.org.

Yours sincerely,

Karen Bennett, Vice President, Environmental Affairs
National Mining Association

National Mining Association

Comments on

The Effects of Mountaintop Mines and Valley Fills on Aquatic Ecosystems of the Central Appalachian Coalfields

1/19/2011

Comments on Scientific Advisory Board (SAB) Draft Report dated 11/8/2010

- EPA is utilizing a scientifically flawed study as the basis for implementing a threshold standard for Conductivity of 500 uS/cm on valley fill projects associated with coal mining operations.
    - The "Pond-Passmore" Study (Pond et al, 2008) used 2 multi-metric indices: the peer reviewed WVSCI, which looked at families of aquatic life, and the draft, un-published genus level index, GLIMPSS, developed by some of the authors of report. The study concluded that the WVSCI underestimated impacts associated with mining related disturbance.
    - Recalculation of the data received through FOIA using the GLIMPSS indices suggested that Ephemeroptera metrics had the lowest value of the six metrics for both impacted and non-impacted sites. As there were two Ephemeroptera metrics in the GLIMPSS index, this biased the GLIMPSS scoring to highlight impacts to the sensitive Ephemeroptera macroinvertebrates, with indices which were "UNTESTED" from the beginning.

- Conductivity is typically a surrogate water quality parameter, and findings should be investigated further to assess the underlying parameters responsible for elevated conductivity AND parameters responsible for poor invertebrate success.

- There are few changes of functional feeding groups at conductivity levels between 2500 and 5000 uS/cm. The most sensitive are filter-collectors, when conductivity exceeds 10,000 uS/cm.

- The distance between Valley Fill sediment ponds and sampling sites had an inverse relationship with conductivity, with conductivity levels decreasing with distance from the sediment ponds. The distance from the valley fill should have been documented and utilized in the assessment to determine if distance was a primary or secondary factor in the conductivity level and in the composition of the aquatic community.

- Valley fills utilizing downstream, on-channel sediment ponds may create a flow regime more reflective of perennial stream/spring systems, with the attendant aquatic community than intermittent/ephemeral streams and their associated communities. Flow regimes should be documented and incorporated into the assessment of the aquatic community composition. In addition, reference sites should have comparable flow regimes to impacted sites, including reference sites with ponding, similar to valley fill sites.

    - Filter-feeding insects tend to be found at higher abundances immediately below surface release impoundments. While there may have been slight changes in percent trophic function below ponds, some literature pointed to increases in abundance, suggesting that general functionality of the community was not altered, even though taxonomic composition changed.

- The SAB Dec. 18, 2008 Advisory on Aquatic Life Water Quality Criteria for Contaminants of Emerging Concern (EPA-SAB-09-007) reiterates the need to perform adequate testing on multiple classes, phylums, and families consistent with EPA's 1985 Guidelines.

- The Pond-Passmore study does not adequately characterize the relationship between conductivity and aquatic life to establish a limit at 500 uS/cm.

- There is an existing regulatory structure to address direct and cumulative impacts from surface mining, mountaintop removal and utilization of valley fills.

    - A PEIS on Mountaintop Mining/Valley Fills was issued in October 2005 by the US Army Corps of Engineers, the US EPA, and US Department of Interior's Office of Surface Mining (OSM), and the US Fish and Wildlife Service (FWS), and the West Virginia Department of Environmental Protection to evaluate options for improving agency programs and communication under the Clean Water Act (CWA) the Surface Mining Control and Reclamation Act (SMCRA), and the Endangered Species Act. The preferred alternative enhances environmental protection and improves efficiency, collaboration, division of labor, benefits to the public and applicants. The document and its associated protocols and legal agreements are designed to inform more environmentally sound decision-making for future permitting of MTM/VF.

- The continued analysis of data collected during implementation of the CWA Section 404 program by the COE and possible future identification of minimal and cumulative impact thresholds has the potential to reduce the number of fills and minimize fill sizes.

- The information sharing and automation of data relative to aquatic resources should also have a positive effect on minimizing fills, individually and cumulatively.

References:

Pond, G. J., M. E. Passmore, F. A. Borsuk, L. Reynolds, and C. J. Rose. 2008. Downstream effects of mountain top coal mining: Comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools. *Journal of the North American Benthological Society* 27: 717-737.

Pond, G. J., J. E. Bailey, and B. Lowman. *The West Virginia GLIMPSS (Genus Level Index of Most Probable Stream Status): A benthic macroinvertebrate index of biotic integrity for West Virginia's wadeable streams*. Unpublished draft report, US Environmental Protection Agency Region 3, Wheeling, WV.

US EPA. 2005. Mountaintop Mining/Valley Fills in Appalachia Final Programmatic Environmental Impact Statement. EPA 9-03-R-05002

Geotechnical
Environmental
Water Resources
Ecological

GEI Consultants

Memo

To: Karen Bennett, National Mining Association

From: Bob Gensemer and Steve Canton

Date: January 19, 2010

Re: Comments on SAB Draft Report dated December 28, 2010 for Quality Review: "Review of Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams"

GEI has reviewed this latest version of the EPA Science Advisory Board's (SAB) comments on the conductivity benchmark report (EPA 2010) for which GEI already presented detailed comments (GEI 2010). In most respects, the SAB report is relatively unchanged from the October version, but they continue to assert that when considering uncertainties in EPA's draft conductivity benchmark of 300 µS/cm, that there are many arguments to suggest that this value may actually be under-protective. While we have commented on this in general previously, we would like to point out several specific SAB statements which we suggest are generally speculative in nature, rather than representing supportable conclusions from the scientific literature or the EPA benchmark document. These statements from section 3.6 of the SAB report (pages 23 – 25) and our responses are summarized below.

- "The field-based benchmark is probably more reflective of how the invertebrate community responds to conductivity than would be chronic toxicity tests. One reason is that chronic toxicity tests usually involve abbreviated times of exposure (relative to generation times of species) and they use surrogate species." The SAB goes on to assert that one of the major flaws in the use of surrogate species such as *Ceriodaphnia* is that they "are not especially sensitive to changes in major ion concentrations…" and that "the species most sensitive to conductivity are often very difficult to work with in demanding tests such as chronic toxicity tests (SAB report, pg. 23)
    - *GEI Response*: The sensitivity of *Ceriodaphnia* to ion toxicity relative to that of mayflies or other presumed "sensitive" species has simply not been adequately tested, and so this assumption is not supported with existing science and requires additional study to confirm or reject. As we pointed out in our report (GEI 2010), the existing literature shows the sensitivity of mayflies to ion toxicity to be highly variable, and often strongly overlapping with the sensitivity of *Ceriodaphnia*. The latter species is widely known for its sensitivity to ion toxicity, so the relative lack of experimental evidence with mayflies to the contrary strongly challenges the SAB's assumption that *Ceriodaphnia* are not particularly sensitive. Also, the SAB is only assuming that any difficulty in conducting chronic toxicity tests or bringing new species into the laboratory are somehow related to that species' sensitivity to major ions. We are not aware of any scientific literature that specifically links ion sensitivity to the difficulty of culturing or testing in the laboratory. In fact, controlling the ion composition of laboratory test solutions is extremely simple, and so can easily be manipulated to suit the sensitivity of any organism.
- "…the survey data have exceptional ecological realism and provide a stronger basis for inferring causality between concentrations of one or more constituent ions (using

GEI Consultants, Inc./Ecological Division
4601 DTC Boulevard, Suite 900, Denver, CO 80237
303.662.0100 fax: 303.662.8757
www.geiconsultants.com

conductivity as a surrogate measure) and presence/absence of genera in aquatic communities in streams below MTM-VF activities." (SAB report, pg. 24)

  - *GEI Response:* We have detailed elsewhere our many concerns with EPA's demonstration of causality for ions related to conductivity being responsible for macroinvertebrate species distributions in the field (GEI 2010). In the above statement, the SAB is suggesting that use of field data alone provides a stronger basis for establishing causality. We suggest that quite the opposite is true, and that it is far simpler and more straightforward to establish causality in the laboratory, or particularly in controlled field manipulations such as mesocosm tests or whole stream tests, which have not yet been conducted in this context. Field survey data are purely correlational, and cannot by themselves be used to establish causality, no matter how compelling the statistical relationships may be.

- A general concern is stated on page 25 that using an SSD to derive a benchmark based on an HC05 is not necessarily protective of the most sensitive species, particularly those with sensitivities (i.e., XC95 values below 300 µS/cm) below the benchmark value. (SAB report, pg. 24)
  - *GEI response*. This is similar to frequently made complaints on most numeric aquatic life criteria derived using EPA methods. As clearly stated in EPA's own guidance for derivation of aquatic life criteria (Stephan et al. 1985), Clean Water Act criteria are not intended to provide protection for all species at all times, but rather to protect the overall integrity of the aquatic community. EPA continues to use this protection level in derivation of aquatic life criteria, and so we see no reason to conclude that the 95% protection level is insufficiently protective to achieve the goals of the Clean Water Act. In addition, the inappropriate use of the term "extirpation concentration" by EPA in their HC05 derivation can easily mislead a reader into thinking actual species will be eliminated if conductivity is increased over 300 µS/cm. In fact, no species are "extirpated" – it is simply an analysis based on the probability of capture, which has many components involved, not just water quality. Extirpation is a process that results from a series of biological factors that lead to the local loss of a taxon from a particular stream or region. To our knowledge, such a process has not yet been demonstrated as a result of elevated ions related to conductivity.
- A general concern is stated on page 26 that a benchmark based only on the distribution of macroinvertebrates may not represent protection levels of other members of the aquatic community, including mussels and fish. Indeed, this leads SAB to state that this problem can "…"thereby make it less defensible."
  - *GEI Response*. We continue to agree with the SAB in this respect, and are surprised that this does not represent a stronger challenge to the appropriateness of the conductivity benchmark, or any other benchmark derived on the basis of a limited portion of the aquatic community.

References

GEI Consultants, Inc. (GEI). 2010. Technical Review: A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams. Report 091380, National Mining Association, Washington, DC. 71 p.

Stephan, C. E., D. I. Mount, D. J. Hansen, J. H. Gentile, G. A. Chapman, and W. A. Brungs. 1985. Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic

Organisms and Their Uses. PB-85-227049. U.S. Environmental Protection Agency, Office of Research and Development, Duluth, MN.

U.S. Environmental Protection Agency (EPA). 2010. A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams. Draft EPA/600/R-10/023A, U.S. Environmental Protection Agency, Washington, DC.