Issue Paper: Considerations for setting a CBP goal and outcome for stream quality

Prepared by Scott Phillips (USGS, chair of STAR CBP Nontidal (NT) Workgroup), Claire Buchanan (ICPRB), Peter Tango (STAR coordinator), Jennifer Greiner (USFWS, Habitat Goal Team Coordinator) and Jeff Horan USFWS, Habitat Goal Team Chair) (updated March 26, 2013)

Question: Should the CBP adopt the Stream Restoration Outcome in the President's Executive Order (EO) strategy as part of the CBP new goals/outcomes or should other options be considered?

The CBP is considering using the EO stream outcome as one of items in our revised goals and outcomes. **The EO Stream Restoration Outcome is:** ***"Improve the health of streams so that 70 percent of sampled streams throughout the Chesapeake watershed rate three, four or five (corresponding to fair, good, or excellent) as measured by the Index of Biotic Integrity, by 2025".*** The EO stream outcome is based on the CBP Stream Health indicator, which was developed during 2006-2008 and approved by the partnership in 2008. The indicator showed that 45% of the sampling sites rated fair, good or excellent, 52% of the sites rated very poor or poor, and 3 percent were not rated.

In 2010, the Federal partners, in consultation with the States and DC, used the information from the CBP stream health indicator to set an EO strategy outcome for 70 % of the streams to be fair or better by 2025. The scientific basis for the outcome was that as practices are implemented in the watershed to reduce nutrients, sediment and other pollutants (such as efforts to reduce toxic contaminants and acid mine drainage), this will provide improvements in the quality of streams. The current target (a benthic index of biotic integrity rating of fair, good or excellent) is also closely tied to the Bay estuary target of meeting water-quality standards for oxygen, clarity and chlorophyll-a in 60 percent of the Bay segments by 2025. We should have a greater percentage of improvements in streams in the watershed by 2025 since they will respond to management actions prior to seeing improvements in the estuary. Therefore a target of 70 percent was chosen (see appendix A for more information on technical rational of the outcome).

Concerns: As the CBP partners consider updated goals and outcomes, there are concerns using the current EO stream outcome as a basis for a new stream goal. The Commonwealth of PA has stated that more scientific analysis is needed to set a steam health outcome because the current indicator has limitations for goal setting (see appendix B). ICPRB, who has worked with UMCES development of the CBP indicator, has responded to the PA comments (appendix C). Additionally, more recent analysis by ICPRB suggests that a technically valid method to assess annual changes in stream conditions cannot be applied to the CBP stream health indicator due to the various state monitoring schedules and approaches. The current CBP stream health indicator may only be able to be updated every 5 to 10 years, which would limit its use as an outcome.

Recommendation: While the CBP stream health indicator is still a very useful tool to quantify the status of stream health and help guide restoration and protection efforts, other options should be pursued to consider a goal/outcome for stream conditions and annual process toward the outcome.

Suggested Next Steps:

- The Habitat Goal Team advocates that a goal/outcome for stream condition should be considered by the CBP partnership. The CBP recognized the importance of stream quality by having commitments to improve stream habitats in the Chesapeake 2000 agreement. Additionally, benefits of Bay water-quality restoration will be seen first in streams so states, local jurisdictions, and citizens groups want to have more information on steam conditions. These streams also support important fish and wildlife species that are being considered as part of the CBP goals (such as Brook Trout). Finally, the 17 million people in the Bay watershed use streams for drinking water and recreation so they are an important way to link people to the restoration of the Bay and its watershed.

- The CBP Habitat Goal Team would set up a working group to assess (a) an improved metric for stream health, and (b) options for an associated stream health goal. The working group would include members of the STAR NT WG workgroup (including ICPRB and Bay jurisdictions that helped develop the CBP stream health indicator), Habitat GIT Stream WG, and other interested parties. Suggested options to for the working group to consider include:
    - Assess if there are other approaches to assess change over time for the current CBP stream health indicator. If so, develop goal options based on this indicator;
    - Develop a stream health metric based on a suite of stream information including fish IBIs, benthic IBIs, stream connectedness to floodplains, and stream bank stability. This will take more effort but may provide a more representative metric and associated options for an outcome; or
    - Have the stream goal/outcome focus on conditions needed to support watershed fisheries. The goal/outcomes could be more closely tied to conditions to support Brook Trout or other key species that will return to streams as fish passage projects are implemented. This option provides less information on overall stream health but does provide a more direct connection to priority fish species in the Bay watershed.

Appendix A: Supporting information in the Executive Order for Stream Condition Outcome

EO Stream Restoration Outcome: ***Improve the health of streams so that 70 percent of sampled streams throughout the Chesapeake watershed rate three, four or five (corresponding to fair, good, or excellent) as measured by the Index of Biotic Integrity, by 2025.***

Current Condition:

45 percent of sampled streams are rated fair, good, or excellent.

Background:

Why is it important? Restoring water quality in streams is a necessary step in meeting water quality standards in the Bay. Similarly, actions to reduce nutrients, sediment and other pollutants flowing into streams to achieve Bay standards also improve the quality of local streams. Restoring streams also benefits the fish, wildlife and people using them. This outcome also helps address comments from the public and states stressing a need to improve local streams as a way of better engaging watershed organizations and involving the 17 million watershed residents in the restoration effort.

What is the measure? This measure of stream quality is an existing CBP indicator, based on an index of biotic integrity which scores benthic macroinvertebrate communities on a scale of poor to excellent. The CBP has worked with the states to gather information from 10,452 sites across the watershed where samples have been collected during 2000-2008 and scored the average of this data based on ecoregion thresholds. For this measure, an acceptable benthic community is defined as having a score of fair, good or excellent. In the future, the CBP will take a subset of these sites to look at change in stream quality over time. The subset of sites will be designed to adequately represent the distribution of stream conditions throughout the watershed.

What is the current condition? In the most recent assessment, conditions at 4,656 sampling sites are rated fair, good or excellent (45 percent), 5,459 sites are rated very poor or poor (52 percent) and 337 (3 percent) are not rated at this time.

What is the basis for the target? The basis for the target is that as practices are implemented in the watershed to reduce nutrients, sediment and other pollutants, we will

see improvements in the quality of streams. The current target (a benthic index of biotic integrity rating of fair, good or excellent) is closely tied to the Bay estuary target of meeting water-quality standards for oxygen, clarity and chlorophyll-a in 60 percent of the Bay segments by 2025. We should have a greater percentage of improvements in streams in the watershed by 2025 since they will respond to management actions prior to seeing improvements in the estuary. Therefore a target of 70 percent was chosen.

More Information:

http://www.chesapeakebay.net/status_watershedhealth.aspx?menuitem=26057

Appendix B: Pennsylvania State Agency comments on using the CBP stream health indicator or goal setting, February 15, 2013
PA DEP COMMENT:
Rodney Kime | Division Chief, Water Quality Standards
Department of Environmental Protection

Summary Comment: We have continually expressed our concerns to the group that developed the "CBP indicator that measures stream quality using an index of biotic integrity which scores benthic macroinvertebrate communities". Simply put, the measure is out of necessity simplistic because it must be applicable to several states encompassing different physiographic regions using different collection methods and the macroinvertebrates are only identified to Family. The indicator is acceptable for use in some limited contexts but we have reservations using it as such an important environmental measure. The measure is not robust and is not sensitive enough.

DEP has spent years in the development and use of macroinvertebrates as indicators of water quality. As a result, DEP biologist are aware of the many confounding factors that must be considered in each stream assessment and that only robust assessment protocols can be trusted. In addition, different protocols are needed in different situations; small freestone streams, limestone streams, limestone influenced streams, pool/glide streams, warm water streams, large warm water streams, and rivers among others. Impairing or de-listing a stream in the context of a 303(d) listing is a serious matter. DEP would never consider using the CBP indicator in such cases. Meeting the Chesapeake Bay restoration targets is an equally serious matter. Just as in the case of 303(d) listings, it would be imprudent to use a generic scaled back Family level IBI to measure progress. DEP reluctantly agreed to the use of the CBP as a broad brush assessment tool but objects to its use as a programmatic tool with broad implications.

One of the major concerns is a reduction in nutrients. The CBP indicator and other macroinvertebrate IBIs are not good indicators of nutrient concentrations. Why would one choose a single indicator especially when that indicator is not responsive to a major pollutant of concern? An EPA Scientific Advisory Board among other things found that:

1. “The Committee emphasizes the importance of choosing the biological endpoints. We note the response of benthic indices can be related to many types of stress. We question why periphyton would not be a better indicator.”

2. “Measurements of actual biological responses would be appropriate, emphasizing variables that respond most directly to changes in nutrient concentrations. These

are typically measures of primary productivity or primary producers, or water chemistry changes such as DO and pH."

3. Moreover, response variables can be at multiple levels - primary response variables (e.g., plants), secondary response variables (e.g., dissolved oxygen [DO], pH), and tertiary response variables (e.g., fish, macroinvertebrates). A change in a response variable is unlikely to be satisfactorily described by changes in a single "causal" variable (e.g., total nitrogen [TN] or total phosphorus [TP]

DEP does not object to using macroinvertebrate IBIs in the context of measuring progress. However, the CBP indicator is not the answer. Neither is relying on a single indicator. Progress should be measured based on multiple important indicators.

PA DEP COMMENT:
Kenn Pattison, P.G. | Hydrogeologist
DEP Water Planning Office

The assumption that 60% of Bay segments in attainment correlates to about 70% of non-tidal streams meeting the IBI category of fair, good or excellent is questionable. The assumption is based on the premise that improvement in IBI scores as result of nutrient and sediment reductions within the non-tidal watersheds is directly proportional to projections of improvement in Bay water quality as a result of the same nutrient and sediment reductions. The efficacy of this assumption is questionable.

First, the large majority of watershed impairments in Pa are the result of acidic mine drainage and sedimentation. The nutrient reductions needed to reach the Chesapeake Bay TMDLs will not result in improvement within those watersheds impacted by acidic mine drainage. As a result, reaching the 70% improved in non-tidal streams would not be reasonable, assuming only nutrient and sediment reductions.

Second, free-flowing freshwater streams are mostly phosphorous limited in Pa. The Bay is a saltwater/brackish water estuarine environment. Once you get beyond the influence of the non-tidal inflows, principally within the middle and lower Bay, the Bay is mostly nitrogen limited. It is unlikely that nutrient and sediment reductions would generate directly proportional water quality improvements within the two completely different ecosystems.

Third, the nutrient and sediment "pollution diet" for the Bay is based on habitat restoration specific to the Bay. Freshwater benthic invertebrates respond to impairments and stressors specific to the freshwater environment. It is unlikely that identical impairments of dissolve oxygen, chlorophyll-a and water clarity standards within the Bay exist in free flowing freshwater streams and that the identical solution of nutrient and sediment reductions would serve to improve the impaired portions of the fresh water streams.

Appendix C: ICPRB responses to Pennsylvania comments on the CBP stream health indicator and goal setting, Prepared by Claire Buchanan, 3/12/2013

Below are responses to some comments made by Rod Kime and Kenn Pattison, which were included in the Pennsylvania state agency comments inserted into the draft

"Chesapeake Bay Program Goals and Outcomes" of January 2013, and in a separate document dated Feb 15, 2013.

1) The measure is out of necessity simplistic because it must be applicable to several states encompassing different physiographic regions using different collection methods and the macroinvertebrates are only identified to Family.

RESPONSE: The measure is not simplistic but rather versatile. Family-level identifications allow the index to be applied to more data sets and geographic areas than it would be if genus-level metrics were used. Some other points: 1) The index follows a common, science-based method for multi-metric index development. It is the same method used to develop the BIBI for Chesapeake tidal waters. 2) The Chesapeake states use biometrics consisting of macroinvertebrates "only identified to family" even when they identify to genus level in the lab. Examples are the Hilsenhoff Biotic Index, %EPT, %Chironomidae, %Diptera, %Ephemeroptera, and %Trichoptera, (some of these are technically order-level metrics). These metrics are used as diagnostic tools or included in the state IBIs. 3) The Chessie BIBI avoids the pitfall of missing—or finding—a statistically rare genus in sample counts (only 100 individuals counted per sample in some states), which can bias taxa richness metrics.

2) The indicator is acceptable for use in some contexts but we have reservations using it as such an important environmental measure.

RESPONSE: We agree that it might not be an appropriate measure for state regulatory purposes since it is not included in state water quality standards. As stated in the May 2011 final report to the CBP Non-Tidal Water Quality Workgroup, "*at this point in its development, the index is useful for regional evaluations and targeting to help managers and watershed groups focus stream restoration and protection efforts. It is not intended for state regulatory purposes and should not circumvent the indexes of stream health developed and used by state monitoring programs for regulatory purposes.*" Although the index is relatively new, it has been tested in a number of analyses and is proving to be sensitive to disturbance. We believe it has strong potential for use as a programmatic measure of stream restoration progress. Although we, the developers of the index, might have chosen a somewhat different goal, goal-setting is the purview of the program leaders. We did not have a say in setting the current goal.

3) The measure is not robust … is not a sensitive measure of the macroinvertebrate community.

RESPONSE: The classification efficiency of the Chessie BIBI—meaning its ability to identify 'reference' and 'degraded' sites correctly—is 91.7% in the Piedmont bioregion, 81.9% in Valleys, 85.5% in Ridges, 86.1% in Northern Appalachian Plateau and Uplands, and 77.8% in North Central Appalachians. These percents are comparable, and even somewhat better, than other indexes developed for the Mid-Atlantic region. One explanation is the larger pool of data used in the analysis. The Chessie BIBI is sensitive to changes in stream habitat and water quality conditions.

4) The CBP indicator and other macroinvertebrate IBIs are not good indicators of nutrient concentrations.

RESPONSE: ICPRB analyses done for Maryland Department of the Environment (Mandel et al. 2011) show that macroinvertebrates are often subject to multiple

stressors as conditions decline in high quality streams, and this masks their responses to nutrients. When samples with confounding stressors such as poor habitat conditions (includes sedimentation & embeddedness), high conductivity, extreme pH, and low DO are removed from the data analysis, a majority of macroinvertebrate (family-level) metrics in the remaining samples show distinct declines as TN and TP increase. The Mandel et al. (2011) study investigated river phytoplankton as well as stream periphyton and macroinvertebrates. It addressed some of the concerns raised in the EPA Scientific Advisory Board review. In another ICPRB study, the Chessie BIBI proved to be sensitive to flow, meaning the index shows statistically significant declines as flow was altered from baseline conditions by land and water uses alteration (ACOE et al. 2012). Of particular impact was impervious surface area in a stream's catchment. In both studies, environmental thresholds for the onset of degradation—different from impairment—were identified.

5) The indicator would miss smaller shifts in the biology.
RESPONSE: Might these smaller shifts in biology be a) artifacts of the 100- and 200-count protocols, or b) indicative of seasonal and geographic differences?

6) The CBP indicator is not specifically designed to measure the effects of nutrients. If the habitat is poor, the CBP indicator will always be low no matter how much the nutrients are reduced.
RESPONSE: The index was designed to be sensitive to the overall quality of the stream, of which nutrient enrichment is one factor. It has turned out to be sensitive to nutrient enrichment, however (see 4 above).

7) A macroinvertebrate IBI by itself is not a good choice [as a biological endpoint] when nutrients are the stressors of concern.
RESPONSE: True. An indicator representing primary producers that directly use nutrients, such as periphyton or biofilms or filamentous algae, is actually a better metric. Basin-wide data to develop this kind of an indicator is not available at this time.

8) Measurable biological impacts due to nutrients occur only under specific conditions. Not all sites will exhibit the impacts.
RESPONSE: This is because of the confounding effects of other stressors (see 4 above). To see an improvement in stream health, the principal stressor would need to be addressed first. Focus on another stressor first and there will be little improvement in biological status.

9) Macroinvertebrates often do not measurably respond to changes over a wide range of nutrient levels, especially using a Family level IBI. Recent to be published experiments using various doses of phosphorous in artificial streams showed no major changes in EPT taxa in doses ranging from 20 to 50 ug/l [equivalent to 0.02 – 0.05 mg/liter] as long as aeration was adequate.
RESPONSE: The figure at the bottom of the next page shows some family-level macroinvertebrate results from Mandel *et al.* (2011). Most family-level metrics, including the Chessie BIBI and %EPT, responded significantly to nutrient enrichment when other environmental stressors were not impacting stream. Thresholds for the onset of degradation in macroinvertebrates were evident at 0.012 – 0.029 mg TP/liter

and 0.58 – 1.13 mg TN/liter, depending on ecoregion. Our phosphorus thresholds for macroinvertebrates are comparable to those in the soon-to-be-published experiments mentioned above.

10) The assumption that 60% of Bay segments in attainment correlates to about 70% of non-tidal streams meeting the IBI category of fair, good or excellent is questionable. The assumption is based on the premise that improvement in IBI scores as result of nutrient and sediment reductions within the non-tidal watersheds is directly proportional to projections of improvement in Bay water quality as a result of the same nutrient and sediment reductions. The efficacy of this assumption is questionable.

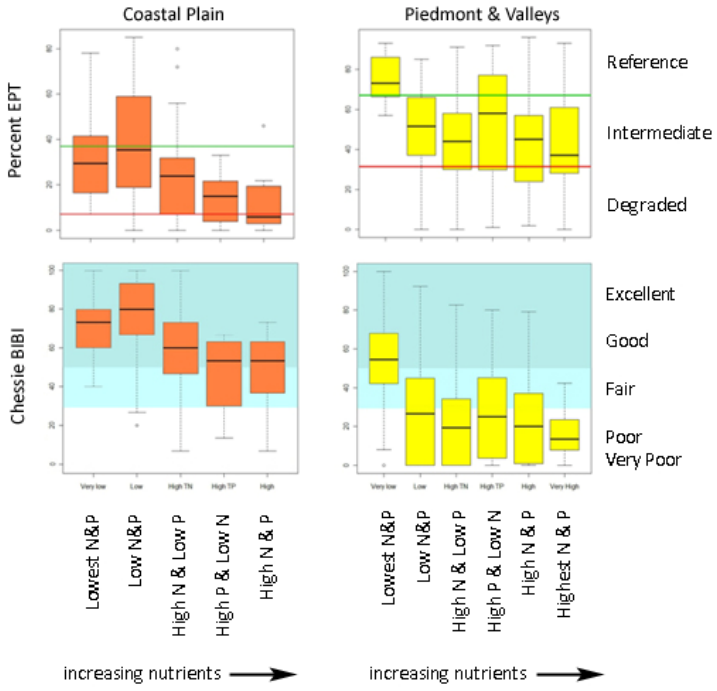
RESPONSE: Agree.

11) The large majority of watershed impairments in PA are the result of acidic mine drainage and sedimentation. The nutrient reductions needed to reach the Chesapeake Bay TMDLs will not result in improvement within those watersheds impacted by acidic mine drainage. As a result, reaching the 70% improved in non-tidal streams would not be reasonable, assuming only nutrient and sediment reductions.

RESPONSE: Agree. Resolving a stream's acid mine drainage problems, however, should increase the stream community's ability to retain and incorporate nutrients into bugs and fish, and reduce the amount of nutrients passed downstream to the Bay.

12) Free-flowing freshwater streams are mostly phosphorous limited in PA. The Bay is a saltwater/brackish water estuarine environment. Once you get beyond the influence of the non-tidal inflows, principally within the middle and lower Bay, the Bay is mostly nitrogen limited.

RESPONSE: Phosphorus limitation does not mean that excess nitrogen is not a problem. Some analyses, including those in Mandel et al. (2011), show that excess nitrogen can actually depress macroinvertebrate metric values in 1st – 4th order non-tidal streams in some ecoregions (see Coastal Plain "High N & Low P" nutrient category in figure). Natural regional characteristics and other factors (e.g. DOC, alkalinity & hardness, karst geology) can and do change the nutrient dynamics in stream ecosystems.



Responses of %EPT and Chessie BIBI to nutrient enrichment when other stressors (e.g., extreme pH, low DO, high conductivity, poor habitat) are not impacting stream. Nutrient categories were determined with a recursive partitioning approach called Category and Regression Tree analysis.

13) It is unlikely that nutrient and sediment reductions would generate directly proportional water quality improvements within the two completely different ecosystems.

RESPONSE: True. CBP models presumably are telling us what nutrient and sediment reduction levels in the watershed will generate the desired improvements in the Bay.