2. Development of Benthic Macroinvertebrate and Algal Biological Condition Gradient Models for California Wadeable Streams, USA. Draft for submission to Ecological Indicators. Paul et al., 2018.

FINAL Compiled Comments
15 March 2019

BCG Product Purpose and Application:

The California State Water Resources Control Board (Water Board) is implementing "…a program to protect biological integrity in wadeable streams using thresholds of numeric [biological] indices to protect narrative aquatic life use (ALU) expectations." The project's goal is to establish policy as water quality objectives that "…protect the biological integrity of wadeable streams from the impacts of eutrophication and other stressors."

The Biological Condition Gradient (BCG) models are being developed to "…support the interpretation of numeric Biological Integrity (BI) goals based on percentiles of reference condition" and to guide decision-makers towards "…more informed decisions about threshold setting." Advantages of the BCG identified in the report include:

- A standardized biological response to a stress gradient
- Universal meaning and application
- Does not vary across regions

In sum, "…a useful construct for interpreting biological indices and for comparing and reconciling regional differences in reference condition, types of indices, or even indices for different assemblages."

The intended use of the BCG models is to "…help decision-makers connect narrative water quality goals to quantitative measures of ecological condition by linking the protective value of targets based on statistical distributions (e.g., percentiles of reference distributions) to narrative descriptions of biological condition (e.g., maintenance of natural structure and function) that are more meaningful to managers and the public."

To support and foster use of the BCG, the stated study purposes or objectives are to:

1. Assign Benthic MacroInvertebrate (BMI) and algal BCG taxon attribute scores for all taxa;
2. Assign BCG levels to 250+ California wadeable stream samples along human disturbance gradients;
3. Develop narratives of structural and functional changes along the BCG gradient specifically associated with degradation of California wadeable streams; and
4. Compare BCG levels [in 6 classification categories] with [California Stream Condition Index for benthic macroinvertebrates] CSCI and [Algal Stream Condition Index] ASCI numeric scores to relate percentile of reference thresholds.

Methods:

The study used two, independent expert panels for macroinvertebrates and, to assign taxa attributes, calibrate the BCG class levels to CSCI and ASCI scores, and conduct and reconcile study-site expert scoring and assignments of biological condition to one of 6 class levels. Level 1 represents the highest biological condition representative of sites in the most natural state and

Level 6 representing the most degraded, or stressed, sites. BCG expert scores were compared to CSCI and ASCI index values, which were developed independently by statistical assessment of biocondition relative to suite of metrics indicative of a stress gradient.

The expert evaluations were designed to calibrate a BCG model to the conditions typical of California wadeable streams. Expert assessments of biocondition were first checked for geographical bias and effects of environmental variables (watershed area, site elevation, temperature, precipitation). Individual stressor indicator variables (TN, TP, conductivity, % agricultural, % urban) were used to form stressor gradients and compared to expert Level assignments to form the BCG. BCG assessments and class level assignments were blind to stress indicators or variables.

Results:

Most taxa were assigned by the expert panels to attribute levels III (Sensitive) and IV (Indiscriminate). Similarly, for both BMI and algae, most class assignments were in the middle levels, 2 – 5, with few assigned to Levels 1 and 6. The relationship between BCG levels and CSCI scores were described as “robust” in the report, and consistent across most regions; ASCI scores also declined with BCG levels, but “not as systematically” and were “more compressed” with substantial overlap in some cases, which may limit discriminatory ability across a full range of biocondition. ASCI was also more inconsistent among regions compared to CSCI.

BCG scores for both BMI and algae were significantly correlated with environmental variables, which were provided to the experts: watershed area (algae an exception), elevation, temperature and precipitation. Correlations to stressors, which were not provided to the experts – TN, TP, specific conductivity and percentages of agricultural and urban land cover – were statistically significant. The authors note, however, that some environmental factor associations might be forced by underlying stressors, e.g, agriculture and urban areas are more prevalent at lower elevations.

Study Conclusions:

- California wadeable stream BCG models showed close agreement among experts within each assemblage
- BCG levels differentiated BMI and algal condition along the disturbance gradient, including:
    - TN and TP
    - Specific Conductivity
    - Agricultural Land Use
    - Urban Land Use

  This may *reflect common stressors…associated with human disturbance gradients of these regions.*
- BCG exhibited “strong” correspondence with CSCI index scores, but “weaker” with ASCI. The authors noted that *This finding is not surprising given that both the indices and BCG level assignments are based on the same comprehensive monitoring program dataset, using a uniform set of protocols, and representing a similar range of conditions.*
- The correspondence between BMI and algal BCG may serve to alleviate problems that are not inherent in expert models, such as:
    - Effect of natural variability, likely to occur in the large, topographically diverse state of California

  - Ambiguity at the index level (attribute and classification assignments)
  - Effect of arbitrary combinations of stressor metrics in statistically-based bioassessment indices
- BMI and algae responded differently to the stressor gradient:
  - BMI BCG categories were broadly distributed over the entire range of CSCI scores, indicating good discrimination potential
  - Algal BCG bins were more compressed and did not extend into attributes 1 and 6 and across the full stressor gradient, i.e., no sites were assigned to BCG levels 1 and 6.

  This disparity was related to, among other reasons, *differential responses of algae versus BMI to stress, more tolerance among algae, incomplete autecological understanding, incomplete range description of least to most disturbed, or scoring bias.*
- Suggested Guidelines for Use of the CA BCG Framework included:
  - It is not intended to substitute or supersede the CSCI or ASCI to interpret taxonomic data, only to provide decision support on the selection of management thresholds (i.e., a percentile of reference)
  - It should not be used to determine biological potential
  - BCG narratives, particularly at the BMI or algal taxa level should not be used to determine probable cause for impaired biological integrity. Causal assessment (e.g.,CADDIS) is recommended for use in such circumstances
- Bottom Lines:
  - The BCG can support decisions for biological integrity goals by providing understanding of the ecological implications of different thresholds
  - BCG levels provide context to the pool of reference sites based on a “best available” definition. If moderately-disturbed sites are part of the reference pool, these reference-based indices potentially assess against a degraded benchmark.

Overview of Science Panel Comments and Suggestions from April 2017 related to BCG

- Consider using a watershed condition index (WCI) as a measure of overall stress in the BCG
- Indices developed over smaller spatial units tend to be more precise and sensitive at detecting stressor responses; this may be a benefit for the BCG applications at finer scales
- If WCI is a good predictor of receiving water biointegrity, especially if the BCG relationship is linear, it may provide a viable conceptual foundation for integrated watershed (IWM) and ecosystem-based management (EBM) applications, with potential for site-specific O/E assessment anywhere along the stressor gradient
- Distinguish between water column nutrient concentration limits, and loading limits for implementation
- Treat discrete point source pollutant loads and diffuse nonpoint/stormwater loads as subsets
- The absence of a quantitative link between beneficial uses and nutrient criteria in the BCG is a concern
- The BCG cannot capture either the inherent natural variability in reference conditions or the sensitivity/vulnerability that different communities exhibit in response to stress
- Reference-condition communities responded differently to the same stress implying that trajectories of response depend on either initial (reference) taxonomic composition or local environmental conditions

- Greater differences among communities are expected with differential responses to types and combinations of stressors
- Reduce confusion by casting alteration as departure from the natural range of conditions expected at a site and acknowledging all assemblages have inherent, quantifiable uncertainty
- Develop a more detailed conceptual model of progressive, beneficial use degradation along a nutrient stressor gradient and relate response to nutrient concentration with probabilities
- Linear responses are sensitive along all increments of the stressor gradient; they are, thus, best for assessing condition responses to stress at any point along the gradient
- A generalized BCG using an ecosystem-level indicator of overall stress paired with a biointegrity indicator that generally reflects overall ecosystem-level stress may be the best way to assess aquatic life designated use and guide EBM to a variety of targets (e.g., natural condition, best attainable condition, threshold/changepoint)
- BCG provides a viable decision-support framework for assessing and targeting management options and actions to specific or interim endpoints anywhere along the gradient, which can support flexible phased and adaptive implementation

Panel General Comments and Suggestions:

The study's technical objectives are appropriate to developing a Biocondition Gradient (BCG) using expert assessments and California data and scoping decision-support applicability to BI goal interpretation using reference percentiles. However, despite meeting the technical objectives of the study, it is unclear how the BCG will be applied and contribute to the programmatic goal of "…using thresholds of numeric [biological] indices to protect narrative aquatic life use (ALU) expectations." The "Guidelines for Use" stated in the Study Conclusions above seem to limit application and miss intended uses suggested in the Practitioner's guide (EPA, 2016), i.e., to "…more precisely describe existing, or baseline, biological condition; help evaluate potential for improvement in condition; and measure incremental changes in condition along a gradient of human disturbance, i.e., anthropogenic stress." Similarly, many of the Science Panel's comments summarized above were not addressed. Consequently, it is unclear how the BCG models will "…help decision-makers connect narrative water quality goals to quantitative measures of ecological condition by linking the protective value of targets based on statistical distributions (e.g., percentiles of reference distributions) to narrative descriptions of biological condition (e.g., maintenance of natural structure and function) that are more meaningful to managers and the public".

One of the identified benefits of a flexible, BCG approach is to support Integrated Watershed Management (IWM) in an Ecosystem-based Management (EBM) framework. A BCG better represents cumulative stress along a gradient rather than "thresholds" in a "Generalized Stress Axis" that captures the effect of multiple stressors on biological integrity indices (*See* BCG and Stressor pathway Figures from EPA (2016) below). Translating a "gradient" effect and relating the technical objectives into desired stressor targets or criteria that meet biointegrity outcomes may require additional work, or at least further explanation.

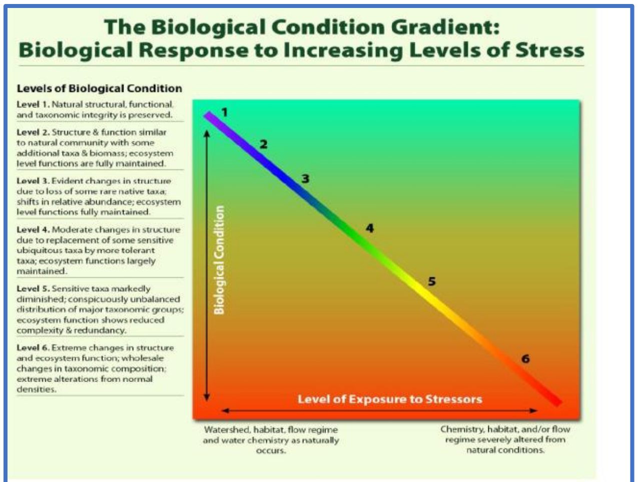


Figure 2. Conceptual model of the BCG. Although in reality the relationship between stressors and their cumulative effects on the biota is likely nonlinear, the relationship is presented as such to illustrate the concept.

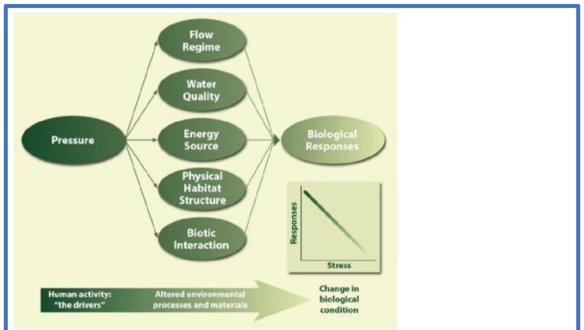
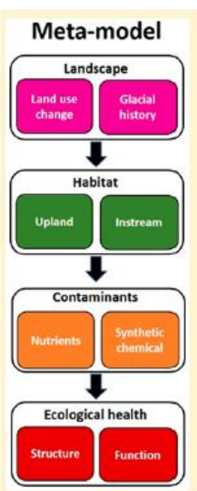


Figure 3. Model illustrating the multiple pathways through which human activities may exert pressure on an aquatic system by altering fundamental environmental processes and materials, creating stressors that may adversely affect the aquatic biota (Source: Modified from figure courtesy of David Allen, University of Michigan).

The relatively poor discriminatory skill of the current CA BCG is one obstacle to applications, especially on a site-specific basis. Discrimination might be improved with a more robust and comprehensive landscape indicator of stress (e.g., combined urban and agricultural percent cover) along the General Stress Axis (x-axis in the BCG). Both the Landscape Model (Beck et al.) and BCG analyses show stronger correlations between biointegrity indicators and percent land cover indicators than those between biointegrity and individual biostimulatory stressors (e.g., TN, TP). The figures below[1] from a Structural Equation Modeling analysis by Schmidt et al. (2019) exemplify complexities of relative stressor effects for three communities representing biointegrity in midwestern streams. Disturbed land cover or its complement, natural land cover, may provide a better indicator of total stress and, thus, a stronger correlation to CSCI and ASCI reflective of responses that theoretically extend across the entire BCG. With a landscape stressor perspective, the BCG could be used to predict by extrapolation class levels 1 and 6 biointegrity scores, for example, levels that had few sites assigned by experts. The meta-model below may provide a more complete, ecosystem-level picture of stressor-response pathways that could be used to improve the current conceptual model.



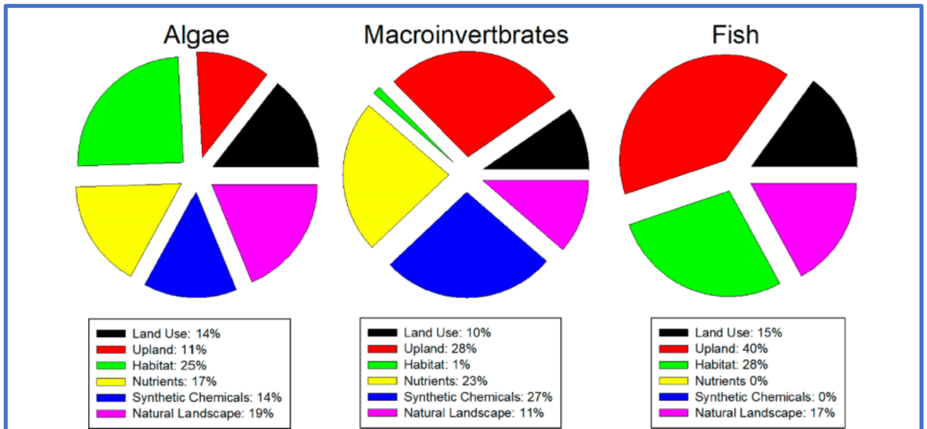


Figure 5. Pie charts illustrating the percent total effects for each class of variable for each of three structural equation models describing the stressors to ecological communities in small streams of the Midwestern United States. See Table S1 for a description of the variables included into each of the six classes described.

[1] Used without permission – for internal project use only, not to be publicly distributed.

The BCG relationship between land cover (e.g., the Watershed Condition Index (WCI) suggested above) and biointegrity would more effectively describe whole ecosystem stress (causal) across the gradient to whole ecosystem response (effect) that would support an EBM framework. EBM would improve understanding and provide a potentially better decision-support foundation for meeting program objectives of "…using thresholds of numeric [biological] indices to protect narrative aquatic life use (ALU) expectations".

If, after considering uncertainty and investigating a more robust land cover indicator of stress, the expert assessments based on the Benthic Macroinvertebrate Index (BMI) and on algae still differ substantially, then it would be interesting to hear from the experts why the two assemblages elicit different responses. Perhaps the assemblages are responding to different stressors, e.g., algal responding more strongly to nutrients and BMI to other stressors or trophic alterations. Additional detail on level assignment uncertainties, and expert suggestions for reducing them, should be reported and may be insightful. A summary would help decision-makers and stakeholders see the proportion of expert disagreement on the assignments and perhaps lead to an integrated BI-algae assemblage model as shown in the figure above.

The CA BCG Framework, with a landscape condition indicator of overall stress suggested above, might better support the intended analytical and assessment foundation to "…more precisely describe existing, or baseline, biological condition; help evaluate potential for improvement in condition; and measure incremental changes in condition along a gradient of human disturbance, i.e., anthropogenic stress" described by EPA (2016). The BCG relationship provides a relative indicator of stress "along a gradient of human disturbance" using readily-available land cover data adaptable to any geography for any existing, historical or goal scenario to be evaluated. It has the capacity to compare predicted historical reference conditions to desired biointegrity goals or endpoints and to scope targets that direct management goals and actions more effectively, a primary intent of BCG applications. Importantly, it is a direct link to a primary IWM/EBM response of landscape management with aquatic life designated use outcomes.

The BCG can also improve individual stressor, including TN and TP, assessments that can facilitate individual stressor targeting in a reference or threshold approach while retaining the ecosystem and gradient logic and perspective. For example, TN and TP loads can be estimated on a site-specific basis by applying export coefficients to reasonably accurate and very precise land cover data. They can be based on historical, current, "best attainable condition" (BAC) or targeted land cover conditions that correlate to biointegrity targets described by the BCG.

Enrichment Factors (EF) should be explored as an alternative indicator of biostimulatory stress of individual nutrients. EFs are ratios of nutrient loads for a current, endpoint, changepoint, BAC, TMDL, Use Attainability Analysis (UAA) or any management scenario of interest, to historical or other reference nutrient load estimate or prediction. For example, an EF of 1 might represent equal reference and current nutrient loads, i.e., no enrichment under current condition. An EF > 1 would indicate nutrient enrichment proportional to the reference.

EFs can also be used as the "stressor" for endpoint or changepoint analyses using TITAN, which also reduces uncertainty caused by natural variability, that are a de facto site-specific nutrient criterion. EFs allow for site-specific, numeric translations of narrative goals that can be used to set permit limits or landscape EBM targets aimed at biointegrity goals and outcomes in a holistic, EBM context. They realistically apportion individual pressure contributions to the impact with no assumptions about other changes in pressure if they are managed individually

(*See* Becker, 2014, for a TP application successfully applied in Connecticut as an alternative to a numeric nutrient criterion).

Care must be taken when applying discrete thresholds, endpoints or changepoints as criteria, even when framed as EFs. Criteria are often set on the brink (i.e., threshold) of environmental change, and management with such precision is not only impracticable, but subject to environmental and response uncertainties. The BCG offers an alternative of working towards a Best Attainable Condition (BAC), reflective of a more natural, historical reference condition, as well as a protocol for documenting incremental and progressive changes in condition towards biointegrity (EPA, 2016), providing a “margin of safety” from the thresholds of change. This flexibility may be preferable for more cautious, or pragmatic, adaptive management approaches in the complex structural and functional ecosystem settings with many uncertainties or unknowns in cause-and-effect relationships.

Response to Specific Charges:

- ***Charge 1:*** Comment on the adequacy of the statewide bioassessment data set and the analytical approaches to evaluate the range of natural variability and its interpretation in CSCI and ASCI.

Data should be better characterized and qualified in the report to answer suitability/adequacy questions. Without that context, data are assumed to be “truth” and error due to sampling (accuracy, completeness, representativeness) is folded into analytical error. However, the study results do suggest that the bioassessment data are adequate to characterize conditions; it should be discussed more fully in the report rather than by reference.

An overview of structural and functional elements of California streams would add perspective. Some biogeographical data were provided to the expert panels that could be expanded upon, and might describe some of the results of the analysis, i.e., attribution to natural variants as opposed to stressors.

The foundation for the taxa traits used in the BCG expert assessment should be discussed in more detail. It may be in citations, and the data seem to reflect those attributions, but some expansion of the foundation might be helpful to the understanding by less well-versed policy makers and stakeholders, as well as being more supportive of the science and its application.

The value of biological data as an integrator of variable conditions over time is an important attribute that supports temporal suitability. More detail on the data, index period and other sampling protocols meant to control variability would be helpful. Spatial density may be limiting ability to capture a full range of responses with adequate discrimination power among forcing factors. Also, since state variables with presumably limited spatial and temporal density are used as indicators of stress, they may not be reflective of the full exposure period of the biological data.

The (stressor) data set holds some concerns. Better characterization and qualification with data quality objectives would assure and communicate adequacy and quality for the intended purpose, i.e, the data should be evaluated for validity as an indicator of process and stress. Driver-level indicators of aggregated stress, e.g., watershed integrity condition indicator, would incorporate a more comprehensive sum of pressures (See Figure 3 from EPA (2016), above) on the aquatic ecosystem, and reduce errors from disentanglements of single stressors from complex structural and functional arrangements. It would likely be more useful to guide IWM/EBM on an watershed scale rather than single pollutant targets or limits.

- ***Charge 2:*** Comment on the adequacy of data set, the analytical approaches and findings of the development of a BCG model.

Methods are generally fine, and consistent with standard BCG protocols. Using logistic regression to estimate the probability of transition between two classes on the BCG (e.g., >= 3 to < 3) would allow a more robust quantification of uncertainty. As noted above in the General Comments, application of the Landscape Model and a predictive configuration might provide more insight into expressions of integrity and their integrated structural and functional causes of stress. Use of an ecosystem-level driver such as watershed ecosystem integrity (e.g., a WCI) based on a land cover indicator, is suggested rather than individual pressures. This seems evident in both the Beck et al., report as well as the results of the BCG analysis presented in Figure 8 of the Paul et al. report. Although it does not look promising, adding a Percent Natural or WCI stressor graph to Figure 8 might show a better relationship to the BCG level distributions or explain why there are so few in levels 1 and 6, i.e., perhaps because there are no waterbodies reflective of those conditions defined by the WCI.

Methods for BCG and Tiered Aquatic Live Use (TALU) application are well developed, but the two are not effectively linked in the report. Since standards and management generally default to the most sensitive designated use, which is almost always ALU support, this relationship is likely a key to effective management applications, and better achieving the intended uses of the BCG described in the study goal on the first page.

- ***Charge 3:*** Are there technical ways to address stakeholder concerns?

Stakeholders presented extensive and useful comments for improving the analysis and its application to setting numeric thresholds to protect biological integrity, and support aquatic life designated use in wadeable streams. Questions of relevance to the BCG ranged from analysis to application. These are hopefully captured in the following themes:

- Application of a general BCG to a numeric water quality “threshold” or endpoint
    - Is this appropriate and supportable as an alternative to water quality threshold or endpoint approaches to protect stream biointegrity?
    - How well can BCG discriminate condition along the gradient and among sites?
    - How does attainment of thresholds or endpoints relate to beneficial use attainment? How does natural variation affect this relationship?
    - How does BCG relate to a Tiered Aquatic Life Use (TALU) application?
    - Can TALU be applied in a phased, incremental or adaptive approach, aimed towards “best attainable condition” of biointegrity?
    - Can BCG models be used to support adaptive management along a BCG “continuum” without using TALU “bins” to set site-specific nutrient endpoints or other indicator targets?
    - How can “best achievable condition” be incorporated?

- Reference condition assessment and application
    - How is “reference” defined and applied in the BCG?
    - Is the reference condition sufficiently defined and distinguishable in the BCG analysis?

    - Is reference condition a consistently definable target for biointegrity?
    - How can natural variability of reference be assessed and placed in context for setting thresholds and management targets?

- Stressor gradient and attribution to pressures
    - Don't stressors other than eutrophication pressures (e.g., nutrient concentrations) impact biointegrity? Aren't causes much more complex?
    - Are eutrophication indicator thresholds adequately correlated to biointegrity indices to guide target-setting and management?
    - How can BCG be related to multiple lines of evidence or a broader stressor index that would consider more than one stressor at a time?
    - How can the BCG model be used in a causal assessment to determine most likely stressors responsible for poor biological condition beyond nutrient effects?
    - Has a predictive distributional approach been considered as an alternative to single threshold values?

- Land cover and biological condition outcomes and constraints in a watershed context
    - Does a watershed integrity indicator consider a broader suite of stressors, including non-biostimulatory stressors?
    - Is land cover a driver of aggregated stress reflected in biointegrity indices?
    - Is land cover an effective indicator of biointegrity constraints?
    - Can land cover be the indicator of collective stress and management targeted to land use condition?
    - Won't a watershed approach account for site-specific conditions, stressors and relationships that can't be adequately captured in the statewide BCG analysis?
    - How can BCG guide narrative objectives for biointegrity when landscape constrains ability to achieve numeric thresholds? [*See also,* Best Achievable Condition]

- Management of biostimulatory stressors and biointegrity outcomes
    - Will biostimulatory stressor management alone control cultural eutrophication?
    - Can reference and target conditions be met?
    - Does the BCG model predict biointegrity targets will be achieved if nutrient targets are achieved?
    - How does BCG and TALU facilitate communication of science and policy?

Most of these issues were addressed or at least touched on in the Comments and Charge responses above. BCG/TALU targets by level (tier) reflect varying levels of biointegrity, which may be framed as policy to accommodate a flexible range of policy and management options to preserve, manage or mitigate reflective of recovery/management potential. It may also be a better framework in application for adaptive, incremental management paced to meeting interim goals/by tier.

Concerns about range vs. single number endpoints could be addressed by using process model applications that better link forcing factors with outcomes, and application on smaller spatial scales that reduce the effect of spatial and site differences. Or, the alternative analytical suggestions for applying the BCG and relating it to a land cover driver may overcome some of the concerns with regard to reference application, stressor – response relationships along the BCG, and site-specific applicability. These differences might not be apparent in a coarser, generalized state-wide assessment, and presume similar outcomes that are unlikely across the board. Becker (2014) is offered for a review using nutrient enrichment factors (EF) to translate individual stressors into numeric management targets, or criteria, on a site-specific basis.

It would help to clarify if BCG is meant as a communication tool for understanding what reference means conceptually or if it is meant to help guide selection of metrics or interpret their response to human-caused stress or both. This seems to be suggested in the Bottom Line conclusions above, as well as a concern raised by stakeholders, but the mechanism for those uses is not specified. Similarly, how the BCG will "…provide decision support on the selection of management thresholds (i.e., a percentile of reference)" needs added explanation. As an alternative to a threshold or changepoint approach with an incremental improvement management strategy, it differs from targeting endpoints or criteria and how TALU might be applied as an alternative to ambient criteria for nutrients through threshold analyses.

References:

Beck, M.W., R.D. Mazor, et al. 2018. Prioritizing management goals for stream biological integrity within the developed landscape context. Draft Manuscript submitted to Freshwater Science. 51 p.

Becker, M.E. 2014. Interim phosphorus reduction strategy for Connecticut freshwater non-tidal waste-receiving rivers and streams technical support document. Connecticut Dept. of Energy and Environmental Protection, Hartford, CT. 70 p. https://www.ct.gov/deep/lib/deep/water/water_quality_standards/p/interimmgntphosstrat_042614.pdf

EPA. 2016. A practitioner's guide to the biological condition gradient: A framework to describe incremental change in aquatic ecosystems. EPA-842-R-16-001. U.S. Environmental Protection Agency, Office of Water, Washington, DC. 227 p.

Schmidt, T.S., P.C. Van Metre and D.M. Carlisle. 2019. Linking the agricultural landscape of the midwest to stream health with structural equation modeling. Environ. Sci. Technol. 53(1):452–462.