NEIWPCC Executive Summary of Peer Review Responses for *Assessment of Nutrient Loading and Eutrophication in Barnegat Bay–Little Egg Harbor, NJ in Support of Nutrient Management Planning*

December 17, 2014

I. Background

Implementation of the Clean Water Act has allowed for significant improvement in the coastal water quality of the New York-New Jersey region, resulting in the restoration of many historical uses and functions of these waters. Despite this success, coastal water quality problems remain and additional actions are needed to fully restore coastal waters. Many waters experience periodic harmful algal blooms and declining trends in submerged aquatic vegetation and desirable fish and shellfish species.

Nutrient over-enrichment of coastal, shallow water embayments in New York and New Jersey leads to eutrophic conditions in those waters. Enriched nutrients in bay waters promote the growth of opportunistic organisms, such as certain species of phytoplankton and macroalgae. Due to the complex relationships between nutrient loads, ambient water column concentrations, and environmental fate and effects, the development of additional or alternative nutrient criteria and management scenarios is needed for these systems.

From 2009-2013, Rutgers University, in collaboration with the U.S. Geological Survey (USGS), executed the *Assessment of Nutrient Loading and Eutrophication in Barnegat Bay–Little Egg Harbor (BB-LEH), NJ in Support of Nutrient Management Planning* project (Nutrient Assessment) for the New England Interstate Water Pollution Control Commission, as part of the *Nutrient Assessment and Management for Shallow Coastal Embayments in New York and New Jersey* grant from the United States Environmental Protection Agency. This interdisciplinary Nutrient Assessment project, conducted by Rutgers researchers Michael J. Kennish, Benjamin M. Fertig, and Richard G. Lathrop and USGS researchers Ronald J. Baker, Christine M. Wieben, and Robert S. Nicholson, integrated models of the coupled BB-LEH watershed-estuary system to estimate levels of nutrient loading and employed a suite of key water quality, biotic, and habitat indicators for quantifying and characterizing estuarine responses and eutrophic conditions associated with these environmental stressors at local and estuary-wide scales.

Due to the complexity of the BB-LEH ecosystem and the comprehensive scope of the project, NEIWPCC elected to request guidance from a Technical Advisory Committee (TAC) comprised of local stakeholders. The group included staff from New Jersey Department of Environmental Protection, the United States Environmental Protection Agency's Region 2 office, and the Barnegat Bay Partnership's Program Director and Program Scientist. The TAC corresponded and met via conference calls and meetings (as often as monthly at some points during the project) to provide

input throughout the project, from the initial stages through the final draft Nutrient Assessment report.

Upon receiving the Nutrient Assessment report, NEIWPCC, in coordination with the TAC, elected to conduct an external peer review with three well-qualified reviewers, nationally recognized experts with broad expertise in estuarine science and statistical analyses. The goal of this peer review was to independently evaluate the methods and findings of the final Nutrient Assessment report and to determine the utility of these methods both for assessing the health of the BB-LEH ecosystem and informing subsequent management decisions. The TAC recommended peer reviewers and then worked to develop the final list of potential reviewers, taking into account potential conflicts of interest. Potential peer reviewers who declined to participate were asked to provide recommendations of other well-qualified experts.

The three participating peer reviewers, listed below, possess a strong mix of knowledge and expertise in estuarine ecosystem science and also in the statistical methodologies employed in the development of indicator thresholds and eutrophication indices:

- **Christopher J. Madden, PhD** – Lead Scientist at South Florida Water Management District;
- **J. Ananda Ranasinghe** – Senior Scientist Emeritus at Southern California Coastal Water Research Project; and
- **Roberto J. Llansó, PhD** – Estuarine and Coastal Ecology Senior Scientist at Versar, Inc.

Note: Peer reviews were conducted in each reviewer's capacity as an independent consultant, and should not be considered products of the affiliated organizations.

A Peer Review Scope of Work was developed based on questions raised from the TAC's review of the Nutrient Assessment report, which was provided to peer reviewers to offer context and to focus their reviews. The Peer Review Scope of Work's key considerations and questions, as well as the responses received from the peer reviewers, are summarized in Section III below. The original comments submitted by the peer reviewers are compiled in a separate document titled *Peer Review Responses for Assessment of Nutrient Loading and Eutrophication in Barnegat Bay–Little Egg Harbor, NJ in Support of Nutrient Management Planning*. Upon receiving the three completed peer reviews, peer reviewers were compensated for their work and recommendations, and the TAC convened to review and discuss key results and common points, and to summarize finding and conclusions reached as a result of the peer review. A summary of the common themes and conclusions from the peer reviewers are addressed in Section II. Comments were not used to make modifications to the Nutrient Assessment report, but rather, in line with the peer review goals stated above, will assist in evaluating the methods and findings of the Nutrient Assessment report and determining the utility of the report's methods for assessing the health of the BB-LEH ecosystem and informing subsequent management decisions.

II. Summary of Peer Reviewer Comments

Peer reviewers all commended the Nutrient Assessment report authors on the amount of effort required in pulling together such an extensive picture of nutrient loading and eutrophication in BB-LEH, in terms of the range of different sets of data compiled, processed, and analyzed. All three peer reviewers agreed that the Nutrient Assessment report authors made significant progress in: (1) assembling applicable and available (albeit limited) data, and (2) improving our understanding how nutrient loading affects the condition of the estuary. However, all three peer reviewers agreed that the lack of BB-LEH data limited the Nutrient Assessment report authors' ability to determine threshold indicators. Comparisons to other estuaries and best professional judgment by the Nutrient Assessment report authors supplemented existing data, but with the methods used, the Nutrient Assessment report can only provide a rough approximation of local thresholds for such a complex system.

While the Nutrient Assessment thresholds will require further experimental and/or statistical validation and potentially revision, the report provides a clear picture of nutrient loading, showing that the highest nutrient loadings are concentrated in the northern part of the watershed, where highest percentages of urban and agricultural land use occur. The USGS Nutrient Loading Assessment was praised by all three peer reviewers for providing a valuable, comprehensive study of the nutrient inputs entering the bay from the developed landscape.

The approach taken by the Nutrient Assessment researchers in developing of the Index of Eutrophication was one item for which the peer reviewers all recommended there be additional explanation and/or justification. The methodology used was novel and, without further justification (i.e., analysis comparing it with more standard methods), may increase uncertainty and limit the ability of BB-LEH managers to come to conclusions and recommend actions based on the outputs of the index. Peer reviewers recommended conducting further experimental and statistical validation, including refinements to the approach/methodology and the incorporation of new variables in developing the index value, as more and other data become available that improve the index's utility for informing management decisions.

The inconsistent availability of data was noted by all three peer reviewers as a key constraint to the Nutrient Assessment report authors' ability to achieve the goals of the project. The Nutrient Assessment report and this peer review provide BB-LEH managers with a valuable resource in analyzing the current state of the estuary. A comprehensive compilation of current data and an analysis of data strengths and deficits will be extremely helpful as a reference when prioritizing BB-LEH monitoring needs. The work is a positive step towards understanding enrichment impacts, defining thresholds, and indicating steps needed for remediation in the Barnegat Bay ecosystem.

III. Questions Posed to Peer Reviewers & Summary of Peer Reviewer Responses

Questions were posed to the peer reviewers through a document entitled "Peer Review Scope of Work." The Scope of Work identified five key topics for peer reviewers' consideration, followed by more specific examples and more detailed questions for each topic. The key topics were:

- Data Sufficiency to Draw Conclusions / Develop Index;
- Selection of Thresholds to Define Condition;
- Derivation of the Index;
- Use of Index; and
- Overall Adequacy of the Report.

The following section includes each question posed and any background context provided to the peer reviewers in italics, followed by a summary of peer reviewer responses in bold text.

<u>Data Sufficiency to Draw Conclusions / Develop Index:</u> There are significant limitations with respect to the data available for this project – examples include a lack of data for some measures across the years considered, limited frequency of data collection (typically quarterly monitoring so that if growing season is selected as stated, there may be only two sampling events), and limited data available for some locations. Given these limitations, can the conclusions drawn and/or the index developed be used for management purposes, developing strategies to target specific levels of nutrients that would be expected to result in support of a healthy ecosystem for this estuary? Are conclusions supported by the data? Specific examples of data limitations are listed below:

1. *Data quality concerns: Were censored values, non-detects, zero values, skewness, outliers handled correctly/adequately?*

   <u>Summary of Peer Reviewer Responses:</u> Generally, yes, especially considering the amount of data that was incorporated from a range of datasets.

2. *Are the written conclusions in line with data presented? Should there be any concerns regarding poor statistical correlations?*

   <u>Summary of Peer Reviewer Responses:</u> The Nutrient Assessment report authors' conclusions were considered a legitimate interpretation of the data presented in most cases – poor statistical correlations are expected with ecological data. However, data gaps, asynchronous data, and poorly correlated data required that the Nutrient Assessment report authors make a number of assumptions in their analyses, and therefore the conclusions presented may not be the only possible interpretation. While the Nutrient Assessment report authors' overall conclusion that BB-LEH is a eutrophic estuary is likely valid, a number of the statements made and trends presented are not sufficiently justified in the report.

3. *The study states that, "The BB-LEH database was analyzed for each segment of the bay, because these segments have been determined to be heterogeneous habitats." If this statement is true, is the determination of one threshold calculation for the entire bay for each indicator appropriate in determining the indicator score for each segment, or should*

the threshold calculations for indicators be defined separately for each segment in order to determine the indicator score for each segment of the Bay?

Summary of Peer Reviewer Responses: In general, sediment size and salinity are known to have effects on organism abundance, and these parameters exhibit well-documented differences along East-West and North-South gradients in the BB-LEH estuary. It is expected, therefore, that different segments would have heterogeneous habits. However, the segmentation was determined prior to the statistical justification being conducted, and was instead based on this knowledge of the bay. A more objective, statistical approach to determining segmentation could have improved index calculations and provided better targeting of management actions.

Regarding the necessity of defining threshold indicators separately for each segment, peer reviewers' opinions differed, suggesting that there is no one optimal approach. One peer reviewer expressed approval of the use of one set of values for all segments; the second believed that only some indicators warranted calculation of separate scores; and the third thought that different threshold indicators should be calculated for each segment.

4. *Do current USGS studies sufficiently capture (identify and estimate) all substantive N and P loads to the bay? If not, please identify additional sources that should be considered.*

Summary of Peer Reviewer Responses: Peer reviewers were generally positive on the USGS Appendix, but some additional Nitrogen (N) and Phosphorous (P) sources were suggested. These recommendations provide important insight into data needs that should be highlighted for future projects:
- Potential import of nutrients through the inlet
- Abiotic sequestration, burial or dissimilatory nitrate reduction to ammonium (DNRA), and denitrification
- Airshed inputs
- More detailed groundwater input calculations

5. *Do the included condition variables include all important parameters of interest regarding the bay's condition? Is it important or useful to have any estimates of microbial loop or secondary production (e.g., if only for the bay's herbivores [clams])?*

Summary of Peer Reviewer Responses: A few additional variables were suggested (listed below), but the coverage of this Nutrient Assessment report was deemed adequate for this point in the process of developing indicators and thresholds. Additional indicators would require decisions on the value provided by each, taking data collection costs into consideration. As with the previous question, these recommendations provide important insight into data needs that should be highlighted for future projects:
- At least one higher trophic level indicator (hard clams, fish, etc.)
- Benthic chlorophyll / infauna
- Phytoplankton community
- Measurement of atmospheric deposition

- Atmospheric and watershed nutrient contributions to biotic uptake, biotic degradation, and sediment processes
- Microbial loop assessments (lower priority)

6. *Given the methodology used to derive a unit-less score for the index, the index assessment for any given year is opportunistic (limited by the data available for a given year) and not deterministic (informed by data from the full suite of prospective relevant factors). As a result, the importance of setting thresholds against which observations are compared to determine the assessment cannot be overstated. As the value for each threshold is one of the most important elements in determining the outcome of applying the index, it is essential that the threshold values be solidly based in science. In other estuary studies, the causal thresholds (for nutrients N and P) were selected based on modeling the relationship between the causal factor and the response variables in the particular waterbody, which is appropriate because the fate and transport of nutrients will vary given the physical/chemical/biological dynamics unique to that water body. Here, the causal thresholds were selected before that modeled relationship has been determined. Does this limit the study's use for management purposes, developing strategies to target specific levels of N and P that would be expected to result in support of a healthy ecosystem for this estuary?*

   <u>Summary of Peer Reviewer Responses:</u> Due to the limited availability of data which hampered the ability to develop sufficient modeling analyses of cause-effect relationships, the Nutrient Assessment report authors needed to draw conclusions based on weight of evidence. This approach taken was appropriate in light of the data available.

   Given the current uncertainty around the stated thresholds, its use for management purposes should be done only with caution. Several years of validation of the framework and additional data gathering will be needed.

<u>Selection of Thresholds to Define Condition:</u> The basis for selecting the thresholds is given as literature, data analysis, best professional judgment (BPJ) or a combination of these factors. Is this a supportable basis for selecting thresholds that would be used to make condition assessments and inform management options designed to effectuate improvement in condition? Specific concerns and questions:

1. *Is there sufficient information within the study report to show that there are enough Barnegat Bay data to determine each of the threshold indicator values? Has the report addressed which indicators relied more heavily on BPJ or literature and should be revisited when more Barnegat Bay specific data for that indicator are available? Is the report detailed and transparent enough such that the reader can reproduce all steps taken to get to the conclusions provided?*

   <u>Summary of Peer Reviewer Responses:</u> All peer reviewers agreed that available BB-LEH information is insufficient to determine threshold indicator values with confidence. Compared to relying solely on BB-LEH data, using data and effects from a wider geographic range can reduce uncertainty. Overall, though, the assessment can only get within a range of the true quantitative characterization with the combination of

literature values, best professional judgment, and local data that was used for this Nutrient Assessment report.

2. *Are you aware of any other significant data/studies that are relevant and should be included or referenced in this study and should have been used to help determine the threshold indicator values? Please explain fully.*

   Summary of Peer Reviewer Responses: This BB-LEH Nutrient Assessment report is on the forefront of its field among studies that look to set estuarine thresholds and assess eutrophication. Other relevant studies were suggested, but not stressed, as each estuary requires its own analysis.

Derivation of the Index The derivation of the index relies on a Principal Component Analysis (PCA) and a series of manipulations involving raw data values and weighted scores. Is the derivation of the index in the manner indicated supportable? Specific concerns:

1. *Determination of index values blends raw scores (comparison of average of raw data to a selected threshold) and weighted scores (square of eigenvector value, considering the factors for which there were data in a given year). Weighted scores simply represent a measure of the variability of the factor, if it is present within a given year. If there are no data, the factor is given no weight. What is the purpose of blending the weighted score with the raw score, and is this a valid approach?*

   Summary of Peer Reviewer Responses: The purpose of the blended score is to incorporate a measure of the variability of a component into the contribution of that component to the overall index – one could identify highly variable factors that might be more worthy of scrutiny in monitoring, and hence given more weight. However, because this is folded into the overall score, it removes the ability for a manager to evaluate variability without decomposing the score. The manager would understand how the components combined into the final score, but would need to report three scores for each index component (raw and weighted scores, and the final score) when transmitting information and making management decisions that utilize the scoring system.

   All peer reviewers agreed that this method requires significant justification (more than what was presented) to warrant its use. While the approach taken is not necessarily incorrect, additional explanation and justification appear needed. In its current state, the ability to make management recommendations based on analyses using this index is limited.

2. *The approach taken in using PCA is not standard and no documentation is presented to justify it. Typically, to develop an index using PCA, the scores of the first few principal components would be examined. If the first eigenvalue (score variance) comprises a large amount of the total variability, then the first principal component might be taken as the index. If weighting the index is desired then the first eigenvalue would be used as a weight. In this report, there do not seem to be any attempts to assess the adequacy of using only the first principal component. Is this approach valid? If not, what argument, further analysis, and documentation would justify this approach?*

<u>Summary of Peer Reviewer Responses:</u> Similar to the peer reviewers' responses to the previous question, it was agreed that the approach taken makes the index less intuitive to interpret, and adequate justification is warranted if employing this new methodology.

3. *The approach taken in this report is to use the squared component of the eigenvector as a multiplicative weight for that component of the index. The justification is that this weight would be the variance of the component. Is this claim correct? If the variables had been standardized to a variance of 1, then there would be some basis for this, although correlations between variables would also have to be considered. The SAS[*] code in the appendices shows that no variance standardization was done during the PCA analysis and it did not appear to have been done before that. Should the use of multiplicative weighting not be justified, as well as this particular weighting method? Do these concerns affect the validity of the index's derivation, and what can be done to address them?*

   <u>Summary of Peer Reviewer Responses:</u> The approach taken by the Nutrient Assessment report authors was novel: as standard practices were not employed, differences in the approach and in the results need to be explained thoroughly, especially when attempting to draw conclusions from the results.

4. *The sole justification for combining the weighted and raw indices is that it integrates the multiple indicators and their variability. The advantage of this approach is not obvious and requires some justification and documentation. Would combining the two indices serve to blur any useful measure, or instead improve it? Do these concerns affect the validity of the index measures?*

   <u>Summary of Peer Reviewer Responses:</u> Peer reviewers commented that the Nutrient Assessment report authors need to demonstrate that adding weighted scores is necessary or that it adds value to the index. The validity of the index is potentially (but not definitively) diminished by the incorporation of the variance component in the index – a systemic analysis comparing the method used to alternatives (more standard methods) is needed.

<u>Use of Index:</u> Objective 5 of this study is "To generate an Index of Eutrophication as a tool to evaluate future conditions using water quality and biotic indicators to assess eutrophication, eutrophic impacts, and overall ecosystem health of the BB-LEH Estuary..."

1. *Does the study report provide enough information on how one can use the Index of Eutrophication to evaluate future conditions using newly acquired water quality data? Is the report detailed and transparent enough such that the reader can reproduce all sets taken to get to the conclusions provided?*

   <u>Summary of Peer Reviewer Responses:</u> Enough of the step by step information is available to create and populate the index. However, more supporting information on how to use this tool in the near term as a management program is ramped up is needed. There is too much need to reference the SAS code in an appendix to find answers rather than having the entire rationale and step-by-step method clearly explained, with supporting documentation, in the body of the text.

[*] NEIWPCC Note: Statistical Analysis System, a software suite developed by SAS Institute for advanced data analytics.

2. *In your opinion, what are the weakest and the strongest aspects of the Eutrophication Index and the Threshold determinations? Please make suggestions on how the weakest parts can be strengthened.*

   Summary of Peer Reviewer Responses:

   Strengths:
   The Nutrient Assessment report authors have successfully vetted the data that they did have, and presented it in a comprehensive analysis. The Nutrient Assessment report can serve as a starting point for management conversations, as an inventory of the scientific knowledge about the bay, and a gap analysis to guide future monitoring priorities. It is the foundation for establishing a robust index supported by a well-designed monitoring plan.

   Weaknesses:
   - Index calculations should either be better justified or replaced with a standard method
   - Further validation/analysis of the index is needed (sensitivity to changes in watershed pressures, understanding how individual components influence the final index)
   - Data availability to develop the index was limited
   - Need a detailed understanding of circulation and residence times (hydrodynamic/water quality model)
   - Airshed modeling or wet and dry deposition data should be captured over several years to understand their significance in the model

3. *Are there any elements missing from the Eutrophication Index which you think need to be included or which would strengthen the tool? Please explain fully.*

   Summary of Peer Reviewer Responses:

   Other indicators:
   - Higher trophic species
   - Benthic mats and infauna/macroinvertebrates
   - Phytoplankton bioassays to develop dose-response curves

   Additional processes that need to be understood:
   - Relationship between seagrass/epiphyte colonization and the components of turbidity
   - Denitrification, dissimilatory nitrate reduction to ammonium (DNRA) and Nitrogen sequestration/burial (need process-oriented measurements)

4. *The Estuaries and Coasts article Mind the Data Gap: Identifying and Assessing Drivers of Changing Eutrophication Condition (Fertig, et al.) identifies grouping the variables into three major categories, one of which is seagrass, to develop an index of eutrophication; however, there are no seagrass data available for the first 15 of the 25 years of data used to develop the index. Thus, can we be confident in using and applying this index?*

   Summary of Peer Reviewer Responses: While the current seagrass component of the index is a good start, additional years of data are needed to make it more robust. Peer reviewers expressed different levels of confidence in using this component in its current

state, but all stressed that continuing data collection could improve it significantly.

5. *Does the approach used here "validate" the developed eutrophication index?*

   <u>Summary of Peer Reviewer Responses:</u> Peer reviewers agreed that further validation should be used to confirm the approach and the specific output of the index, and to improve confidence in drawing conclusions using the index. An important example mentioned by one peer reviewer was the need for a stronger analysis and discussion of how changes in Total Phosphorous and Total Nitrogen loadings affect index results – within a single year, across years, and separated by segment.

   The need to further validate this index is one of the many compelling reasons for developing and maintaining a comprehensive monitoring program for BB-LEH. Values may need to be tweaked and additional variables considered for incorporation.

<u>Overall Adequacy of the Report</u>

1. *Is the organization of the document appropriate and does it present the material in a clear and concise manner? Please explain fully.*

   <u>Summary of Peer Reviewer Responses:</u> The Nutrient Assessment report is a positive step towards understanding enrichment impacts, defining thresholds, and indicating steps needed for remediation in BB-LEH. That being said, some assumptions are not fully explained, and there are organizational aspects that make it difficult to follow.

2. *Are the stated objectives adequately met? Please explain fully.*

   [NEIWPCC Note: The seven key objectives of this study are listed below for reference.

   1. *To document the influence of human altered land use on past and present nutrient export from the BB-LEH Watershed to the BB-LEH Estuary using physical and chemical watershed data and land-use patterns, and spatially explicit models.*
   2. *To determine if nutrient loading quantified by subwatershed and biotic response is stable or is temporally and spatially variable.*
   3. *To quantify baseflow, runoff, and total nutrient loads and to determine the relative importance of turf area coverage.*
   4. *To determine estuarine biotic responses to the loading of nutrients across a gradient of upland watershed development and associated estuarine nitrogen loading, and to identify key biotic responses across a variety of estuarine organisms by examining shifts in phytoplankton, benthic macroalgae, seagrass, epiphytes, benthic invertebrates, and shellfish structure and function. Each of these parameters will be examined and assessed for statistical validity and inclusion in the index development for the 1989 to 2010 period.*
   5. *To generate an Index of Eutrophication as a tool to evaluate future conditions using water quality and biotic indicators to assess eutrophication, eutrophic impacts, and overall ecosystem health of the BB-LEH Estuary and to develop threshold levels of biotic decline and numeric loading criteria that can support an effective nutrient management plan.*
   6. *To apply a conceptual model of eutrophication and determine if ecosystem structure and function have been altered in the BB-LEH Estuary.*

7. *To document the current biotic and seagrass habitat conditions of the BB-LEH estuary at the end of the investigation using the most recent biotic data collected (2011) and index methods developed from data collected through 2010.]*

<u>Summary of Peer Reviewer Responses:</u> All peer reviewers generally agreed that Objectives 1-3 were successfully met. One peer reviewer believed that Objectives 4-7 were also adequately met, while the other two thought that these objectives were partially met. The availability of data was noted by all three peer reviewers as a key limitation in the Nutrient Assessment report authors' ability to meet the objectives of this project.

3. *Do the results from the study support the authors' conclusions and recommendations?*

<u>Summary of Peer Reviewer Responses:</u> The results do somewhat support the Nutrient Assessment report authors' overall conclusions and recommendations, especially with regards to the sources of enrichment occurring due to land use changes. Biogeochemical processing within the estuary still requires further study, and further justification and analysis regarding the procedures used in the development of the index is needed. Peer reviewers were not all completely confident in using the Nutrient Assessment report to inform management decisions without first conducting additional validation and refinement. However, all peer reviewers agreed that Nutrient Assessment report authors have an excellent knowledge of the BB-LEH system, and this Nutrient Assessment report is a commendable step in the right direction.