Compilation of Preliminary Comments from Individual Panel Members

(As of June 16, 2016)

Dr. Merryl Alber 1
Dr. James Ammerman 6
Dr. Steven Bartell 8
Dr. Celia Chen 14
Dr. John Connolly 18
Dr. Robert Diaz 21
Dr. Richard Di Giulio 23
Mr. Douglas Endicott 25
Mr. James Fitzpatrick 33
Dr. Robert Heath 36
Dr. J. Val Klump 41
Dr. Douglas McLaughlin 47
Dr. Emma Rosi-Marshall 49
Dr. Eric P. Smith 52
Dr. William Stubblefield 57

Dr. Merryl Alber

EPA Science Advisory Board Lake Erie Phosphorus Objectives Review Panel
Review of Phosphorus Load Reduction Targets for Lake Erie (June 2016)
Initial Comments from: Dr. Merryl Alber, Professor, Dept. of Marine Sciences, Univ. of Georgia Submitted June 15, 2016

Background: The EPA has requested the Science Advisory Board to review the current modeling results and other information used to inform development of binational phosphorus reduction targets for Lake Erie. Below I provide my initial responses to the six questions posed in the panel charge, followed by my comments and suggestions for future work.

Question 1: Evaluation of the models to inform the interpretation of results.

The documents did a good job laying out the technical criteria that were used to identify and evaluate the various models used in the assessment, and clearly explained the process by which load-response curves were generated and compared. The ensemble modeling approach seemed to work well, although there were some compromises because they were not all calibrated and validated against the same observations and there was not time to run all of the models with identical inputs and response data.

1. Some of the models in the ensemble have been calibrated/validated against old conditions. The 9-box model, in particular, provides predictions that are often different from the other models in the suite. For future modeling runs, it would be useful to use standardized inputs and validation data. It will also be important to agree on standardized timing for assessing critical TP loads, predicting summer chlorophyll a, etc., which currently varies among the different models.

2. Some of the mean absolute relative errors (MARE) presented in the document approach 50% and even higher. It would be helpful to have some context for these values as well as the other metrics that were provided (RMSE, PBIAS). How do they compare with other models (i.e. those used in the Gulf of Mexico or Chesapeake Bay)?

3. The results of the U-M GLERL model suggest that the relationship between spring Maumee River TP load and HABs has changed over time, with an increase in blooms for the same load. The updated phosphorus mass balance model found that the same TP load resulted in lower TP concentrations (although this was over a different time scale). This brings up several questions:

> a) Do the NOAA/HAB model or WLEEM models predict the same increase in sensitivity between 2008 and 2013/14 seen using the U-M GLERL model?
>
> b) For a given TP load, are there differences between 2008 and 2013 in other response variables (i.e. TP concentration, overall phytoplankton biomass or hypoxia)?

c) If there is evidence for temporal changes in any of these responses, can the models be used to identify/separate out the potential causes (i.e. changes in % DRP or N:P ratios, the effects of mussels, stratification strength)?

4. The seasonality of both loading and lake response may be shifting. The analysis in Stow et al. (2015) suggests not only increases in March TP and DRP conc and loads over time, but also documents changes in other months and constituents (i.e. decreases in DRP load from June-Sep; decreases in both N and P concentrations in July; increases in P conc and load in December). It will be important to evaluate how these changes affect the proportion of the annual P load that is discharged in different seasons, and whether this has implications for identifying the critical spring loading period used for predicting HAB blooms (now defined as March-July). There may also be potential changes in temperature/stratification strength that might affect which months exhibit hypoxic conditions (which is now being evaluated in August-Sept).

Question 2: Recommended phosphorus loading targets for Lake Erie.

The proposed phosphorus loading targets are generally supported by the modeling and loading information that was provided in the documents, and address the Lake Ecosystem Objectives with respect to minimizing hypoxic zones and preventing harmful algal blooms.

1. Based on the modeling results for 2008 conditions, a 40% reduction in spring loading for the WB appears appropriate to reduce HAB blooms for most years. As described above, however, it will be useful to evaluate whether there has been an increase in sensitivity over time and to recognize that the critical spring period may change in the future.

2. If I am interpreting things correctly, the various models predict that an annual load of 6000 MT to the WB + CB will result in summer (Aug-Sep) average hypolimnetic DO concentrations that range from 2.2-3.5 mg/L. I am basing this on the text on pages 28 in the "Recommended Phosphorus Loading Targets" report, which talks about WB + CB loads and refers to Figure 18. However, the x-axis in Figure 18 is labelled "Annual **whole** TP load" as opposed to "WB + CB", which I believe is incorrect. The text on p. 35 describes it as WB + CB annual load (which I am assuming is correct), although it incorrectly refers readers to Figure 19. This is confusing and the text should be revised.

3. The documents provided describe the rationale for targeting an Aug-Sep hypolimnetic oxygen concentrations of 2 mg/L or more as opposed to 4 mg/L (which would require a greater reduction in load). However, assuming that comment #2 is correct, it should be made clear that an annual WB + CB load of 6000 MT will result in a) a hypoxic area that is larger than 2000 km^2 and b) a greater number of hypoxic days than would be predicted if the objective were to maintain DO at greater than 4 mg/L.

The results from the various models suggest that the target load for achieving a hypoxic area of 2000 km^2 ranges from 2600 to 5100 MT. According to Figure 17 of the "Recommended Phosphorus Loading Targets", at a load of 6000 MT the predicted hypoxic area ranges from approximately 2000 to nearly 6000 km^2.

The models predict that at a load of 6000 MT the number of hypoxic days will vary from approximately 35-40 days, as shown in Figure 16"6" in the "Annex 4 Ensemble Modeling report". However, note that this is another example where the x-axis indicates "whole lake" load (although see Figure 4 in Scavia et al., in press, which is labelled WB+CB). If whole lake load is indeed accurate the predicted number of hypoxic days would need to be adjusted downward given that the WB+CB load is only 88% of the total. This again should be revised in the text.

The average DO concentration, extent of hypoxic area, and number of hypoxic days are all factors that bear monitoring going forward to ensure that the targeted loads are resulting in the expected responses. As described for question #1, it will also be useful to evaluate whether the "summer" response period changes over time.

4. The target of 6000 MT represents a reduction not only in nonpoint source loading from tributaries but also the Detroit WWTP and from atmospheric deposition, so all of these should be taken into account and quantified.

Question 3: Addressing Cladophora growth.

Based on the information presented, I agree that it would be difficult to develop scientifically-sound recommendations to address Cladophora growth at this time and that additional research is needed. However, if the initial extrapolations are correct then it may be that the current loading reductions will be protective.

1. Remote sensing of Cladophora is already being conducted and is a good way to get a synoptic picture of nuisance algae and to track expansion of habitat. It would be useful to evaluate whether remote sensing can also be used to estimate tissue phosphorus content.

2. The modeling that is currently being conducted will be helpful for addressing both nearshore and offshore sources of nutrients, which is information that will be necessary in order to develop appropriate loading targets.

Question 4: Consideration of nitrogen control.

The current focus on phosphorus is warranted given the fact that the Lake responded to the decrease in inputs in the past.

1. Stow et al. (2015) documents a decrease in TN load from the Maumee River since 2000, despite concurrent increases in discharge. They also provide evidence for decreased inputs in summer months (May-July) in recent years, and seasonal shifts in the TN:TP ratio (decrease in Mar-Apr; increase in Sep-Nov). It would be useful to evaluate the implications of these changes with respect to cyanobacterial blooms as well as the other response indicators. If NH_4 is also being monitored it would be interesting to evaluate trends in that as well, as NH_4 it is more available to phytoplankton than NO_3.

2. It would be helpful to develop loading budgets for TN similar to those that are presented for TP (i.e. Fig. 2 and Figs. 5-13 in "Recommended Phosphorus Loading Targets"). This would help answer questions regarding changes in N loading over time (i.e. were decreases due to changes in agricultural sources) and to determine whether N loading is similar to P in terms of the relative importance of the Detroit R. Finally, it would be interesting to calculate the FWMC of NO_3.

3. Nitrogen cycling is likely influenced by the presence of Dreissenid mussels (e.g. Svenningsen et al. 2012) and this may in turn affect N:P stoichiometry and availability for phytoplankton and macroalgae. One way to assess this would be to evaluate denitrification in bottom waters using N_2:Argon ratios (e.g. Grantz et al. 2012).

Question 5: Accounting for inter-annual variability in hydrology in assessing reduction of tributary loadings

Flow-weighted mean concentrations (FWMC) provide a way to calculate the average concentration that enters the lake from a tributary over a given period of time. For example, the FWMC for Maumee spring TP load for 2008 is 0.23 mg P/l. This will be useful to have, although FWMC is calculated after the fact and so cannot be directly compared to a nutrient standard: a given concentration might or might not be a problem depending on the flow.

1. It would be helpful to do an analysis based on long-term flow records to come up with guidance on target FWMC concentrations for the various percentile flows (i.e. 10-90%) for the Maumee and other tributaries. These could be calculated for spring loads (Mar-Jul) as well as for other intervals in order to provide a way to assess whether targets are being met.

Question 6: Evaluation of nutrient reduction goals for Lake Erie

One approach to adaptive management is to modify or test hypotheses iteratively by adjusting management operations. This type of adaptive management will be important for Lake Erie, given uncertainties in the effectiveness of the present phosphorus reduction targets with respect to the response indicators coupled with the potential for changing future conditions.

1. Continuing measurement of TP and DRP loads to the Lake as well as monitoring the response variables (HAB blooms; DO concentrations; Cladophera) will be critical for determining whether load targets are being met and whether they are having the expected response. Key variables that are necessary for both calibrating the various models and evaluating potentially important drivers should also be monitored. This might include measurements of DRP, Dreissenid mussels, and oceanographic conditions. A full list of such variables should be developed, using the recommendations included in the background documents as a starting point (e.g. "Recommended Phosphorus Loading Targets", pp. 45-46.) As part of this it will be important to consider the appropriate spatial and temporal scales for measurements.

2. The change in HAB susceptibility needs to be pursued and better understood. Likewise, changes in other response variables should be evaluated (i.e. for the same TP load, are

chlorophyll concentrations or DO different than they were previously)? It may be useful to choose a later year with discharge similar to 2008 for comparison.

3. There is a need to consider the effects of climate and other factors (e.g. invasive species) on the response variables. Long-term changes in precipitation and discharge can affect the distribution of nutrients, residence time and stratification strength (Stow et al. 2015). Changes in temperature can affect stratification, rates of biogeochemical cycling, and the seasonality of inputs. Additional changes such as introduced species may affect the Lake in ways that are as yet unforeseen (e.g. Asian carp eats Cladophora). Efforts should continue to understand these interactions and determine whether there are identifiable triggers or thresholds that result in large changes in response variables. It may also be necessary to develop additional response variables if new symptoms of eutrophication are manifest.

4. Several of the responses above identified research topics that should be considered going forward (e.g. assessment of denitrification). Other potential topics include:

> a. sediment oxygen demand – the sensitivity analyses for the models suggested that SOD and the link to hypoxia is an important research gap. There is a need to quantify how changes in external loads affect SOD, and whether sediment storage and internal recycling might result in a delayed response to reduced loads.
>
> b. Cladophora – The Cladophora working group identified the need for a coordinated assessment strategy and outlined research and modeling priorities. There is a need to collect data to parameterize models with a nearshore focus that can be used to forecast how P affects biomass in the Eastern Basin of Lake Erie.
>
> c. Evaluating the efficacy of different BMPs with respect to DRP and N will be important.

References:

Stow et al. 2015. Long-term and seasonal trend decomposition of Maumee River nutrient inputs to Western Lake Erie. Environmental Science and Technology 49: 3392-3400.
Grantz, E.M., A. Kogo and J. Thad Scott. 2012. Partitioning whole-lake denitrification using in situ dinitrogen gas accumulation and intact core experiments. Limnology and Oceanography 57: 925-935

Svenningsen et al. 2012. Shell biofilm nitrification and gut denitrification contribute to emission of nitrous oxide by the invasive freshwater mussel Dreissena polymorphja (zebra mussel). Applied and environmental microbiology. 78: 4505-4509

Dr. James Ammerman

Charge Questions:

1. Please comment on whether the evaluation of the models was adequate to inform how model results should be interpreted, given differences in model complexity and scale. Please identify any additional analyses that may be needed to improve future development and interpretation of the load-response curves for the eutrophication response indicators.

I believe that the model evaluation was adequate, though I am sure further comparison tests could be done. As the models continue to develop, some will likely become more complex, but evaluation metrics should also improve.

2. Please comment on whether the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives defined in the GLWQA (as reflected in Table 1 on page 7 of the document titled *Recommended Phosphorus Loading Targets for Lake Erie*).

The recommended targets for TP and DRP appear reasonable, though a 40% reduction will be an extremely challenging target. In the future perhaps some more focus should be put on the NRP, since it is likely to be all bioavailable and may have higher concentrations that the DRP. I realize that the information on NRP is limited, but it is likely an important source of readily bioavailable P and may also vary with the source of P.
NRP

3. Please comment on whether scientifically-sound phosphorus load reduction recommendations to address *Cladophora* growth in the Eastern Basin of Lake Erie could be developed at this time.

Since only one model was presented which addressed this issue, it appears that a limited amount of work has been done. Monitoring for *Cladophora* is also limited at the optimal measurement criteria are still being defined. For instance, there is no standard measure for biomass. *Cladophora* is becoming an increasing problem despite the fact that it occurs in the most oliogotrophic portion of Lake Erie. Further discuss of this issue by the SAB should be useful.

4. What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?

There are good reasons to address the importance of nitrogen control, though perhaps not yet to provide a reduction target. The toxic cyanobacterium *Microcystis*, the major concern in western Lake Erie, does not fix nitrogen and so requires a fixed nitrogen source. It can become nitrogen-limited in late summer in western Lake Erie. In addition, it becomes more toxic under nitrogen-replete conditions. Furthermore, *Microcystis* is very good at scavenging low levels of phosphate, including using alkaline phosphatase to hydrolyze phosphate esters, and thrives offshore in lower phosphate environments. While nitrate is the predominant form of nitrogen in Lake Erie and is highly mobile, there are also lower levels of ammonium and other reduced nitrogen compounds. Since these can be readily used by most cyanobacteria, they could be significant contributors to blooms, even at low concentrations.

5. Please comment on the use of FWMC and any other approaches that should be considered to account for inter-annual variability in hydrology in assessing progress in reducing tributary loadings of phosphorus to the Lake.

Yes, I think FWMC is a useful approach to account for inter-annual variability but should be no means be the sole approach used.

6. Please comment on the value of applying the existing eutrophication models on an ongoing basis to periodically evaluate phosphorus loading targets and eutrophication response indicators. What key elements should be included in the adaptive management approach to successfully implement and evaluate our nutrient reduction goals for Lake Erie?

Continued monitoring and adaptive management is essential to the restoration of Lake Erie, particularly as demonstrated by its history of re-eutrophication. Using a suite of models is usually a good strategy, and the suite to be used will change over time as new models appear, others are improved, and some fade away. Response indicators may also need to change; toxin levels are a response indicator that should be considered in the future. While information may be currently limited, toxin levels are a direct measure of a harmful response indicator and may yield more useful information than cyanobacterial chlorophyll a. Since some cyanobacteria are not toxic, and toxin levels vary with different strains, using toxin concentration as an indicator may improve predictability of adverse impacts. Toxin concentrations may also help to further elucidate the role of nitrogen.

Dr. Steven Bartell

Charge Questions:

Approach for Developing Lake Erie Phosphorus Load Reduction Targets

Nine different Lake Erie models were used to predict the response of selected eutrophication response indicators to different phosphorus load scenarios (see Table 1 in the Annex 4 Ensemble Modeling Report). The eutrophication response indicators evaluated were (1) overall phytoplankton biomass represented by chlorophyll a, (2) cyanobacteria blooms in the Western Basin, (3) hypoxia in the hypolimnion of the Central Basin, and (4) *Cladophora* in the nearshore areas of the Eastern Basin. Technical evaluation criteria were used to assess the capabilities of each model (see Section 2.3 and Appendix B of the Annex 4 Ensemble Modeling Report) and load-response curves were generated for each eutrophication response indicator (see Section 3 and Appendix B of the Annex 4 Ensemble Modeling Report).

1. *Please comment on whether the evaluation of the models was adequate to inform how model results should be interpreted, given differences in model complexity and scale. Please identify any additional analyses that may be needed to improve future development and interpretation of the load-response curves for the eutrophication response indicators.*

It was unfortunate that a more comprehensive evaluation of model performance (using the RMSE, PBIAS, and MARE statistics) was not possible. The value of a "holistic evaluation" is difficult to determine when load reductions are suggested in absolute terms on the basis of uncertain model results, particularly for models that are difficult to compare based on the limited ability to compute the selected model performance statistics.

The individual model presentations in the appendices to the Annex 4 document provide additional statistics for model evaluation and many plots of model results and data. The nature and level of effort in evaluation varies substantially among the models. Taken together, the amount of information is somewhat overwhelming. Perhaps additional effort might be directed at useful summarization of model performance and bringing more of these model evaluations into the main presentation of the document.

It is understood that the models differ in scale and complexity and were not originally developed to define loadings. However, why not at least run the models for the same years to facilitate comparison of model results or at least summarize results for an identical time period for more direct comparisons? Also, it is not clear why all three performance statistics are not presented for an individual model when one is (e.g., Table 2). All three performance measures use the same inputs, y_{obs} and y_{sim}. It is also not clear what is meant by metrics computed or obtained from the literature.

The challenges and difficulties in performing rigorous comparisons and evaluations across the different models is recognized and appreciated. At the same time, it remains difficult to interpret

the value of "visual" comparisons between model results and observations or qualitative statements such as "a good fit", "falls within acceptable ranges", "accurately captures", "approximately the same accuracy in all basins", and "all indicate acceptable performance" in relation to using model results to define actionable decisions that will have real (measurable) economic and environmental consequences.

Normalizing model results to permit across-model comparisons on a percentage scale can provide misleading results to the extent that the models differ in absolute terms (e.g., Figure 10). Even with such normalization, the four models identified in Figure 10 suggest a range of ~3000 MT of TP loading that correspond to 50% of the modeled summer average chlorophyll-a. The appearance of model convergence at higher loadings (e.g., 6,000-8,000 MT) is likely an artifact of the normalization. The unscaled model results should be presented wherever possible, even if this requires log-scaled plots.

Where possible, why not match the averaging periods (e.g., July-August) for model predictions of chlorophyll-a instead of using different averaging periods (e.g., June-August for ELCOM-CAEDYM, EcoLE and July-September for WLEEM)? What was the nature of communications and coordination among the modeling team during the course of the model implementation for generating the load curves for the selected ERIs?

Rigorous sensitivity analysis can provide useful information for evaluating model realism and performance, as well as permit comparisons of alternative models. However, sensitivity analyses of the models described in this assessment were informal and different enough in methodology (ranging from professional judgment, to OAT and Monte Carlo simulation) among the models to make model evaluations and comparisons difficult. In some instances, the sensitivity results served mainly to underscore the original model objectives. For example, the EcoLE was consistently sensitive to parameters associated with mussels (density, grazing rate) – not surprising for a model designed to assess dreissenid impacts on Lake Erie. Encouragingly, most of the models were sensitive to algal growth rates and phosphorus half-saturation constants for algal growth – hence the models appear justifiable for addressing P loading to Lake Erie.

A reasonable path forward in using the models to characterize the likely responses of the ERIs to managed P loading lies in the proposed adaptive management program as outlined in the "Recommended Loading…" document (e.g., Figure 22). The proposed monitoring program provides the opportunity to compare model results with data and correspondingly revise and refine the models for continued application with (hopefully) increased accuracy and reliability. Importantly, the continued model applications and evaluations might indicate an opportunity to focus on a smaller number of models. Alternatively, the strengths of individual models might be used to construct a single model that addresses all of the ERIs and provides a focal point for management and decision-making.

An alternative approach to model evaluation in support of decision-making would be to work with decision-makers to identify acceptable quantitative measures of model performance, including criteria for model scale (number of years, spatial domain) and resolution (daily, seasonal, annual) for the ERIs. The models would then be rigorously and consistently tested

against these decision criteria. This might entail some modifications to the models to permit fair comparisons. One outcome might be that all models fail the criteria. Decision-makers and modelers could examine how relaxing the decision criteria subsequently allowed the various models to make the cut. Decision-makers could also include the costs/consequences of making an incorrect decision on inaccurate model results as part of the overall model evaluation process. Depending on the magnitude of the costs or consequences of possible decisions (e.g., load reductions) and model performance, the decision-makers might recommend additional model refinements. The value and implementation of a multi-model approach should be driven as much by the context and process of decision-making as by the state-of-the-science of eutrophication modeling.

The document, Recommended Phosphorus Loading Targets for Lake Erie describes the process followed by the Annex 4 Objectives and Targets Task Team to develop phosphorus loading targets for Lake Erie. The document indicates that, to achieve a Western Basin cyanobacteria bloom biomass threshold no greater than that observed in 2004 or 2012, 90% of the time, a spring Maumee River load of 860 metric tons of total phosphorus and 186 metric tons of dissolved reactive phosphorus is recommended. In addition, a 40% reduction in the spring load of total phosphorus and dissolved reactive phosphorus from other Western Basin tributaries and the Thames River is recommended. To meet a threshold of 2.0 mg/L or higher of hypolimnetic dissolved oxygen, an annual total phosphorus load of 6,000 metric tons to the Western and Central Basins is recommended. The Task Team did not recommend new phosphorus concentration objectives for the open waters or the nearshore be identified at this time.

2. *Please comment on whether the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives defined in the GLWQA (as reflected in Table 1 on page 7 of the document titled Recommended Phosphorus Loading Targets for Lake Erie).*

If one considers the variability (95% CI) associated with the model predictions for the two empirical relationships and the WLEEM (e.g., Figures 14-16, "Recommended Loadings…"), the targets for cyanobacteria growth could be defined considerably lower, perhaps 500-600 metric tons of total P. The empirical models essentially reflect the best available data to date that describe algal biomass in relation to P loading. The WLEEM appears as a credible process-level representation of algal growth dynamics in the Western Basin. The empirical models represent the drivers of cyanobacteria growth in aggregate form almost by definition. The growth processes described in Appendix B-7 ("Annex 4 Ensemble…") for the WLEEM are substantial in detail and characteristic of models previously developed to simulate phytoplankton production in relation to key environmental factors (e.g., nutrient availability, light, temperature).

Figure 17 ("Recommended Loadings…") illustrates the variability in annual total P loads that correspond to the 2000 km^2 ERI for Central Basin hypoxia. The recommended loading value of 6,000 metric tons appears as an upper estimate associated with an EcoLE result. The recommended loading could be as low as 1000 metric tons (i.e., the nine-box model result) or in the range of 3000-5000 metric tons based on the collective model results. It's convenient that the

6000 metric ton value corresponds to the 40% annual loading reduction suggested for cyanobacteria growth.

In selecting ERIs, why not use maximum 30-day average chlorophyll-a concentrations to help characterize bloom conditions in addition to summer average chlorophyll-a as a measure of phytoplankton biomass? Why only focus on cyanobacteria biomass as indicator of "bloom" conditions? Clearly, cyanobacteria are major contributors to algal production, but do not constitute the entire algal assemblage in Lake Erie.

Cladophora Growth

Additional phosphorus load reductions may be necessary to reduce nuisance levels of *Cladophora* in the nearshore waters of the Eastern Basin of Lake Erie. The Annex 4 Objectives and Targets Task team did not recommend a specific phosphorus objective or loading target to address *Cladophor*a growth. EPA and Environment and Climate Change Canada convened a workshop in January 2016 to assess the current state of knowledge of *Cladophora* growth in the Great Lakes and identify potential options for nutrient target development to be considered by the Annex 4 subcommittee. (Please see the background document titled "State of the Knowledge of *Cladophora* in the Great Lakes. Executive summary of Workshop held at NOAA-Great Lakes Environmental Research laboratory, January 26-28, 2016.")

3. *Please comment on whether scientifically-sound phosphorus load reduction recommendations to address Cladophora growth in the Eastern Basin of Lake Erie could be developed at this time.*

The presentation in Appendix B-9, Great Lakes Cladophora Model (Auer et al. 2016), makes a compelling argument that suggests an initial Lake Erie SRP load reduction of 25% for the Eastern Basin. Despite the incomplete science and uncertainties associated with estimating NAB, as described in the January 2016 workshop summary and underscored in the "Recommended Loadings…" document, the GLM appears no less credible in addressing Cladophora growth than the various models used to address other ERIs in the Annex 4 Modeling Report. Auer et al. (2016) describe a defensible, calibrated process model of Cladophora biomass and internal P concentrations. The sensitivities of the model to input parameter values have been quantitatively described. The model results have been logically extrapolated through a series of load-related models to arrive at the 25% target reduction. The additional research, monitoring, and modeling outlined in the workshop summary are clearly of value. Nevertheless, the resulting target reduction indicated by the Auer et al. (2106) modeling study might be implemented within an adaptive management framework, while waiting for the recommended additional Cladophora work to be completed.

Nitrogen Control

While the current strategy focuses on limiting phosphorus loading to the Lake (total and dissolved forms) as the key mechanism for controlling excessive algal growth, it is implied or assumed that nitrogen loading likely will also be reduced through implementation of agricultural

best management practices, and the Task Team recommended that tributary nitrogen loads to the Lake be tracked in addition to phosphorus.

4. *What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?*

The model descriptions in the Annex 4 Ensemble document suggest that among the eight models (not including the Cladophora model), only the EcoLE, WLEEM, and ELCOM-CAEDYM include state variables for nitrogen. None of the models appear to address internal accumulations of N and P by phytoplankton and corresponding N:P ratios, which could be used to explore possible loading scenarios indicative of N limitation. One or all three of these models might be modified to simulate internal N and P concentrations in phytoplankton. However, it is recognized that many factors (e.g., light, temperature, and other micronutrients) can influence the dynamics of internal N:P ratios in phytoplankton. So, these kinds of modifications would require substantial additional parameter estimates and process formulations. The associated uncertainties and paucity of N:P data for Lake Erie phytoplankton might prohibit these kinds of model modifications. Or, the model results for phytoplankton N:P might be too uncertain to provide useful insights concerning possible N limitation. Nevertheless, such modifications might at least be considered.

Measurements of phytoplankton N:P ratios might be included as part of the detailed analyses performed on the proposed 5-y interval for intensive monitoring. Again, the spatial-temporal variability in phytoplankton N:P ratios and the importance of other environmental factors (in addition to N availability) in influencing these ratios might make it exceedingly difficult to demonstrate N-limitation based on field measures. But it might be worth exploring, if practical.

Evaluation of Nutrient Reduction Targets

The inter-annual loading trends for the Maumee River are greatly influenced by annual variability in flows. The Objectives and Targets Task Team identified a maximum flow below which the target load should be met and recommended the use of flow-weighted mean concentrations (FWMC) as a benchmark for any given tributary target load.

5. *Please comment on the use of FWMC and any other approaches that should be considered to account for inter-annual variability in hydrology in assessing progress in reducing tributary loadings of phosphorus to the Lake.*

The FWMC provides a convenient metric for comparing variable loadings from the Maumee River, either annually or during the “critical period” (March-July). While the metric is convenient for comparing loads in relation to varying hydrologic regimes, the system does not “see” the FWMC. This could have implications for reducing loads. For example, P reductions might result from management actions that reduce P loading primarily outside the critical period.

The corresponding annual FWMC would not capture this and a similar reduction in P loading during the critical period could produce a similar annual FWMC, but the system response (e.g., cyanobacteria growth) to each temporal pattern of P loading could be quite different.

It might be more straightforward to present the plotted hydrograph for the Maumee with a corresponding annual total P (and/or DRP) loading estimate. Or perhaps, flow-weighted monthly concentrations could be computed to provide the temporal resolution helpful in translating loads to ERI responses. These monthly values could be accompanied by the annual flow-weighted concentration to provide for easy comparison of year-to-year variations.

The Task Team recommended development of a comprehensive adaptive management program that would include annual routine monitoring of appropriate load, FWMC, and in-lake nutrient-eutrophication response indicators in conjunction with an intensive monitoring, research, and operational model application program every five years.

6. *Please comment on the value of applying the existing eutrophication models on an ongoing basis to periodically evaluate phosphorus loading targets and eutrophication response indicators. What key elements should be included in the adaptive management approach to successfully implement and evaluate our nutrient reduction goals for Lake Erie?*

It is encouraging that an adaptive management program for nutrient reduction in Lake Erie provides a mechanism whereby results of monitoring will be used to continually revise and refine selected (not necessarily all) models. The results of continually refined and updated models could then be used to help design effective and economical research and monitoring efforts as part of an adaptive management program (as suggested by the adaptive management conceptual model, Figure 22 in the "Recommended Loadings…" document).

An adaptive management program for Lake Erie ought to include the flexibility to quickly undertake intensive monitoring, research, and modeling when monitored conditions strongly suggest that significant (e.g., impaired public use of Lake Erie water) blooms are imminent, based on best available science. Intensive work scheduled at 5-y intervals might appear sensible from an institutional and budgeting perspective. However, this 5-y interval might prove incompatible in scale with the hydrological and biogeochemical processes that influence nuisance algal growth in Lake Erie.

Beyond the scope and authority of this current modeling and P loading assessment effort, any credible and useful adaptive management program for addressing eutrophication in Lake Erie should consider sustainable urban, industrial, and agricultural carrying capacity for all watersheds within the Lake Erie Basin. Technological efficiency gains in controlling point (and perhaps non-point) sources of nutrient loads will eventually be overwhelmed by continued (unsustainable) economic growth. Realistic and effective (feasible) restrictions in nutrient loadings can only be developed within this larger regional economic management context. The best of intentions in managing nutrient loading to Lake Erie will ultimately result in continued eutrophication, if undertaken within an economic paradigm of continuous growth.

Dr. Celia Chen

Celia Chen Comments and Responses to Charge Questions:

First, I would say that one of the most difficult problems for me in reviewing the documents was the inconsistent naming of the different models. This was the case in the earlier report in 2014 and is particularly true in this current Annex 4 report. It makes reviewing the document for those of us less familiar with each model unnecessarily difficult. I would suggest using a consistent nomenclature for the models throughout all reports.

1. *Please comment on whether the evaluation of the models was adequate to inform how model results should be interpreted, given differences in model complexity and scale. Please identify any additional analyses that may be needed to improve future development and interpretation of the load-response curves for the eutrophication response indicators.*

   In general, I found the use and evaluation of the suite of models to be very effective and the consistency of the results to provide strong predictions for the development of recommendations. My only other suggestion would be to use trend data for the increases in precipitation over time to cast forward the impacts of increased precipitation on the proposed targeted P loads and the resulting TP and DRP concentrations. The likely possibility of continuing increases in precipitation with climate change could make the proposed targets ineffective in the future and it would be good to know how far out that future might be.
   .

2. *Please comment on whether the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives defined in the GLWQA (as reflected in Table 1 on page 7 of the document titled Recommended Phosphorus Loading Targets for Lake Erie).*

   I do not necessarily agree that the hypoxia objective has been completely addressed. Minimization of hypoxic zones is based on level >2.0 mg/L even though 4.0 mg/L is the level known to insure enough oxygen for fish and also the level above which hypoxia does not occur. It is not clear to me why the level of 2.0 was chosen. The report states the reason that the Central Basin morphometry and variability in physical conditions. It would be helpful to have a more detailed explanation.

   I am also not convinced that the commitment to protect human and ecological health from harmful toxin levels has been met. While the target loads seem adequate to reduce cyanobacterial blooms, they do not insure that the cyanobacterial biomass levels that would be maintained would not produce toxin levels that would pose threats to human or ecosystem health. The targets are based on estimates of cyanobacterial biomass not on toxin concentrations. It isn't clear to me that the relationship of toxin concentrations to cyanobacteria biomass are known well enough for target loads to insure safe toxin levels.

While I think that the P concentrations to be expected from the loading targets are good, the concentrations for the central basin are not mesotrophic but rather oligotrophic and so the expected concentrations do not entirely meet the LEO of a mesotrophic central basin. In addition, I found the discussion of trophic status in the "Recommended Loading Targets of Lake Erie" report to be lacking in detail (p. 30). That surprised me since earlier discussions focused more on chlorophyll concentrations and there were no graphs to even show the relationships between TP and chlorophyll in the report even though the GLWQA Commitments for open water and near shore waters listed in Table 1 refer to maintaining levels of algal biomass below the level constituting a nuisance condition. The report does not address what a "nuisance condition" would be.

Finally, I am confused about the nearshore objectives as contrasted from the open water objectives both of which are focused on P load. Nevertheless, in the report, the authors state that the nearshore objectives should not be viewed in the context of "re-eutrophication endpoints" since influence of ecological factors is much greater in shallower zones. The report recommends against P concentration objectives for the nearshore areas and recommends for monitoring, research, and modeling for setting nutrient objectives in a broader ecological framework. While this is a good recommendation, it is not clear to me from the report how this will be accomplished.

3. *Please comment on whether scientifically-sound phosphorus load reduction recommendations to address Cladophora growth in the Eastern Basin of Lake Erie could be developed at this time.*

   I agree with the report authors that at present, there is not enough scientific information on the controllers of *Cladophora* growth. Better spatial surveys and method development on determining both spatial coverage and thickness. In addition, more research needs to be done to determine the links between *Dreissenid* populations and *Cladophora*.

4. *What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?*

   I feel that it is important to understand the nitrogen:phosphorus relationships because past literature has indicated that the species composition of Harmful Algal Blooms (HABs) is dependent on both N and P; e.g. limitation of N can cause the switch from *Microcystis* to *Anabaena*. Thus, if N concentrations are increasing, the persistence of *Microcystis* blooms could increase even if P concentrations are lowered. In addition, both inorganic and organic nitrogen species should be distinguished. Davis et al. 2010 found that growth of the toxic *Microcystis* strains were enhance by inorganic nitrogen whereas the non-toxic strains were stimulated by organic nitrogen. Moreover, Zhang et al. 2015 found that microcystin production appeared to be regulated by total N and NO_3 but not by NO_2 or NH_4.

5. *Please comment on the use of FWMC and any other approaches that should be considered to account for inter-annual variability in hydrology in assessing progress in reducing tributary loadings of phosphorus to the Lake.*

   I agree with the use of FWMC in accounting for the inter-annual variability in hydrology. However, I am wondering whether there is variability in the relationship of the discharge volume and the P concentration of the discharge. In other words, without actual measurements of concentrations, how do we know that the calculation of FWMC change doesn't change with volume of precipitation? For example, it would seem that not just the volume of discharge determines the amount of nutrient loading. Larger precipitation events would likely carry larger P loads on a per volume basis. Are there established relationships for this? Having some idea of the influence on tributary input volume and the P concentrations would help to insure that the relationship between hydrology and P concentration is captured properly in estimations of tributary loadings.

6. *Please comment on the value of applying the existing eutrophication models on an ongoing basis to periodically evaluate phosphorus loading targets and eutrophication response indicators. What key elements should be included in the adaptive management approach to successfully implement and evaluate our nutrient reduction goals for Lake Erie?*

   There is great value in the application of the existing eutrophication models to evaluate loading targets and ERIs on a periodic basis. The models need to be tested with newly collected data to determine if they continue to be relevant to the conditions of Lake Erie and to evaluate whether their predictions are sound.

   Among the key elements that should be considered in an adaptive management strategy are the following:

   Relationship of P loading data to TP and DRP concentrations in each basin; and relationship of FWMC to total load target.

   Contribution of resuspended nutrient load from sediments.

   Temporal sampling of FWMC to make sure that year to year variability doesn't alter the frequency and volume of higher level storm and precipitation events.

   Continued measurement of the ERIs along with concentrations and loads to make sure that the predicted relationships do not change.

   Need to include ecological endpoints of interest like fish health, benthic organisms in hypoxic zone, nearshore communities.

   Measurement of N and N:P ratios.

Changes in *Dreissenid* populations and potential influence on P and DRP concentrations in the Central and Eastern Basins.

Changes in coverage and density of *Cladophora* in the Eastern Basin.

Measure of health (particularly Lake Whitefish) and diversity of fish communities.

Dr. John Connolly

Pre-Meeting Comments of John Connolly Pertaining to the SAB Charge Questions:

1. *Please comment on whether the evaluation of the models was adequate to inform how model results should be interpreted, given differences in model complexity and scale. Please identify any additional analyses that may be needed to improve future development and interpretation of the load-response curves for the eutrophication response indicators.*

The modeling evaluation criteria are adequate for judging the applicability of each model to the task at hand and the strength and weaknesses relative to predictions of water quality responses to phosphorus loading changes. However, they seem to have been loosely applied and models were accepted despite widely differing performance against the criteria. The evaluations themselves were hampered by differences in the extent to which each of the models was calibrated, validated and examined for parameter sensitivity. A significant challenge to the evaluation was the assessment of the uncertainty around estimates of loadings at which the water quality thresholds would be achieved. I believe this aspect of the work could be significantly bolstered to have greater confidence that the chosen targets will achieve the desired outcomes.

A gap in the evaluation was an assessment of the consistency and reasonableness of model parameterization. A noteworthy example is the parameterization of sediment oxygen demand. The 9-Box model used a value much lower than the other models and clearly outside the range of measured values cited to in the other model documentation. This apparent miss assignment of SOD was not addressed, but could call in to question the use of this model for evaluating Central Basin hypoxia. There is no consideration of why EcoLE was calibrated and "confirmed" to 1997-1999 conditions using an SOD max of 0.27, which is also outside the range of measured values. The 1-D Central Basin Hypoxia Model adjusts SOD from a base value of 0.75 g/m^2-d to lower values at lower phosphorus loads using a relationship that may be the result of an erroneous interpretation of an equation presented in Borsuk et al. 2001 in which carbon loading to the water column (primary production + external load) is mistakenly replaced with carbon deposited to the sediment. If my interpretation is wrong and the SOD analysis is correct, perhaps the documentation can be enhanced to describe the translation from that presented in Borsuk et al. to that used in the 1-D model.

Also noteworthy is the 1-D Central Basin Hypoxia Model's use of two loading estimates to the Central Basin from the Western Basin, one that decrements the Western Basin load by a factor computed from an assumed net settling velocity and the other from the predictions of WLEEM. The text discusses these in terms of TP, but presumably this was also done for carbon loading. Scavia and DePinto accept the results obtained with the two loading estimates based on concluding "… the two methods yield similar results." (page 26). I took the differences between the methods as possibly significant. The decremented load version computes DO about 30% higher than the WLEEM version. Assuming that the calibrated WLEEM, which fits the Western Basin TP concentrations reasonably well and has calibrated transport to the Central Basin, is superior to the simple model of a constant settling velocity, its greater loading to the Central Basin and lower computed DO seem to be better estimates.

<u>Western Basin Cyanobacteria Biomass</u>

A goal of a bloom no greater than occurred in 2004 and 2012 (nominally a 30-average peak biomass of 1,900 MT) 90% of the time was estimated to be achieved at a spring Total Phosphorus load of 860 MT from the Maumee River. This load reflects load-response estimates derived from a 2008 baseline condition and various approaches to assessing model uncertainty.

Statistical estimates of uncertainty around regression models are used, but these may be inadequate and significantly underestimate the true uncertainty. This is so for several reasons: 1) the overall uncertainty over the full range of loadings may not reflect the uncertainty at loadings in the vicinity of the target; 2) the overall uncertainty may mask trends that produce differing load-response characteristics under different hydro-meteorological conditions or mask trends in time.

It would be instructive to look at the load-responses under different conditions. This was done for the Obenour model, but the finding of a time component was not considered. Both the Stumpf and WLEEM models show a wide range of responses. As seen in Figure B1-3 of Scavia and DePinto, widely varying load responses are obtained, a number of which would indicate that the bloom threshold would not be achieved at the target loading. Similarly, Figure B7-27 of Scavia and DePinto shows that the load-response for two of the four years to which the model was applied indicate the same failure. These results seem to suggest that the bloom would not be less than the threshold goal 90% of the time. These also seems to be borne out by the fact that the 2012 bloom occurred at a loading less than half the target. Perhaps a more thorough examination of the range of predictions by these models coupled with an assessment of the uncertainty of each could be used to get a better sense of the likelihood of meeting the desired goal at the chosen target.

<u>Central Basin Hypoxia</u>

The critical evaluation of the models regarding Central Basin hypoxia pertains to computed hypolimnion DO in the August-September period, the metric used to set the TP target load. It is noteworthy that ELCOM was accepted despite there being no model-data comparisons for DO and no model sensitivity analysis. Wouldn't these deficiencies relative to the evaluation criteria have disqualified ELCOM as a basis for assessing the relationship of TP load and DO? Similarly, it seems puzzling that the 9-Box model would be used for Central Basin DO when its sensitivity analysis showed a lack of sensitivity to SOD, which Scavia and DePinto note to be the most important of the DO sinks, and its use of an SOD outside the range of measurements.

The 1-D Central Basin Hypoxia Model was run for 19 years of hydro-meteorological conditions (1987-2005) and produced a range of hypolimnion DO values at each tested TP load. The mean of these was used in choosing a loading target. It would have been helpful to have presented the frequency at which the DO target of 2 mg/L would be achieved at the TP target loading.

2. *Please comment on whether the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives defined in the GLWQA (as reflected in Table 1 on page 7 of the document titled Recommended Phosphorus Loading Targets for Lake Erie).*

The models differ in their characterization of the drivers of cyanobacteria growth and seasonal hypoxia. For example, WLEEM accounts for many of the factors driving cyanobacteria growth, including selective temperature optimum, buoyancy ability, susceptibility to predation, nutrient uptake dynamics and hydrodynamics, whereas the other models used to set the targets are statistical correlations that take no account of these driving factors. The models used to set targets to reduce hypoxia also differ in their

characterization of the drivers, particularly the extent to which SOD drives hypoxia. As I discuss in response to Charge Question 1, these differences call into question the applicability of some of the models for setting the targets.

The models that best reflect the available information on drivers and seem to meet the test of being well calibrated and validated, which may be WLEEM and the 1-D Central Basin Hypoxia Model, could perhaps be considered as tests of the question of whether the targets reflect the best available information. Under that construct, the answer is yes.

3. *Please comment on whether scientifically-sound phosphorus load reduction recommendations to address Cladophora growth in the Eastern Basin of Lake Erie could be developed at this time.*

I agree with the decision to not develop a recommendation to address *Cladophora* growth. The available model is inadequate to the task given its simplicity and lack of calibration.

4. *What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?*

It seems that the first step is to understand the extent of nitrogen control that will result from mitigation resulting from the TP targets. Additionally, it seems important to understand how nutrient recycling within the lake affects N to P ratios, particularly as loads are reduced to meet the targets. To the extent there are indications of nitrogen limitation and a desire to achieve further reductions in biomass and hypoxia, research to better understand the controlling processes and to develop models that characterize these processes seems warranted.

5. *Please comment on the use of FWMC and any other approaches that should be considered to account for inter-annual variability in hydrology in assessing progress in reducing tributary loadings of phosphorus to the Lake.*

Given the large inter-annual variability in TP loading, it does seem that FWMC is a useful statistic to assess progress in reducing loadings. However, it is also subject to variability and a multi-year record is needed to get a reasonable understanding of the progress being achieved. Another statistic that might prove useful is mean spring high flow concentrations.

6. *Please comment on the value of applying the existing eutrophication models on an ongoing basis to periodically evaluate phosphorus loading targets and eutrophication response indicators. What key elements should be included in the adaptive management approach to successfully implement and evaluate our nutrient reduction goals for Lake Erie?*

Reapplying the same models seems to not provide additional information. However, continued model development and calibration to conditions under reduced loadings could provide useful. The models have considerable uncertainty in key processes like SOD, zebra mussel filtration and nutrient recycling that will be affected by the TP reductions. Continued development seems necessary to address these uncertainties.

Dr. Robert Diaz

Charge Question 4. What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?

Nutrient inputs from point and non-point sources deliver both nitrogen and phosphorus to lake waters. Freshwater primary production is general viewed as controlled by phosphorus availability and that excessive phosphorus loads lead to eutrophication. In Lake Erie there appears to be a causal relationship between tributary phosphorus loading, and both algal blooms and central basin hypoxia. For growth, however, organisms need a combination of nutrients in varying ratios and concentrations. There is evidence for nitrogen co-limitation with phosphorus seasonally and spatially. While the most important nutrient to control is phosphorus, consideration for the role of nitrogen in enhancing eutrophication does need to be considered. Excess phosphorus and nitrogen have set in motion a cascading set of consequences that will not be easy to reverse, but problems created by both nutrients need to assessed.

While not completely similar to Lake Erie, the efforts to control nutrient entering Baltic Sea waters, which also has problems with large blooms of cyanobacteria, have followed a path similar to Lake Erie. A few decades ago the principle nutrient strategy was focused solely on phosphorus. After much research and discussion, the Baltic now follows a dual nutrient reduction strategy including both phosphorus and nitrogen.

There are three components to understanding the contribution of nitrogen to Lake Erie's eutrophication and the degree to which it needs to be control:

1. Total loadings entering the lake temporally and spatially, and the species of nitrogen.
2. The degree to which nitrogen will be reduced or increased in consort with controls for phosphorus.
3. The importance of concentrations and ratios of nitrogen to other nutrients in directing or controlling ecosystem functions, such as nutrient cycling or primary production.

Before a two nutrient reduction strategy is applied to Lake Erie several key issues need to be addressed:

1. How reliable are the current models for assessing nitrogen's role in Lake eutrophication?
2. What balance in the ratio of nitrogen to phosphorus would be best for ecosystem functioning?
3. What would be the expected response in nitrogen reduction within the four groups of management identified eutrophication response indicators?

4. Given the difficulty and expense of controlling and reducing nitrogen loadings what sources and input locations for nitrogen would be most cost effective from monetary and ecological points of view?
5. What can be learned and applied from a case study of nutrient reduction strategies for the Baltic Sea?
6. What are the downstream consequences to not following a dual nutrient strategy?

Dr. Richard Di Giulio

Comments by Richard Di Giulio

Overall I was very impressed by the thoroughness, scientific rigor, and clarity of the two key documents provided for SAB review (Annex 4 Ensemble Modeling Report and Recommended Phosphorous Loading Targets for Lake Erie). The supporting documents (Scavia et al., in review, and State of Knowledge of *Cladophora* in the Great Lakes Workshop) were helpful in better understanding the issues this SAB has been asked to address. Below are my initial comments, made by an environmental toxicologist with very limited modeling expertise.

Charge Question 1: Evaluation of models to inform the interpretation of results.

The discussions in both documents of the nine models evaluated appeared very thorough and objective. The tabular summaries provided indicating which models were applicable to which of the four Response Indicators (Overall phytoplankton biomass, Western basin cyanobacteria blooms, Central basin hypoxia, and Eastern basin *Cladophora*) were helpful. From three to six models were applicable to for all indicators, except only one was applicable for *Cladophora*. As a non-modeler, two questions arose upon reading the material: 1. Where more than one model was applicable, did any particular model appear most helpful? Related to this, on p. 17 of the Annex 4 report, it is noted that annual TP loads to achieve a 50% decrease in chl-a concentrations in the Western Basin ranged among the models between 1130 and 3010 MT. This appears to be a substantial range and I was left to wonder if the modelers had greater faith in one extreme or the other, or neither, and why. 2. "Ensemble modeling" suggests to me that a goal was to develop a model that incorporated selected aspects from those evaluated. But I saw no evidence of this. Are the bases, input data etc. so different that this is not feasible?

Charge Question 2: Recommended phosphorous load targets for Lake Erie.

These targets appear well-justified based upon available data and models, at least as starting points. The development of distinct targets among Western, Central and Eastern inputs, each tied to a distinct Response Indicator, is well-supported (note issue for *Cladophora*, below). It is also important that the proposed targets are in the context of adaptive management. As noted in the reports, the biology and hydrology of Lake Erie, and other Great Lakes, are likely to continue to change with changes in climate, land use, phosphorous chemistry (dissolved versus particulate), community structures (e.g., benthic bivalves) etc. Unfortunately, some changes (for example, warming) are likely to exacerbate eutrophication in the system. Thus, continued monitoring with advanced approaches as they become available will be critical for the achievement of the ultimate goals of this effort. An important complication here is that the lag time between achieving a P loading target and indicator responses may be substantial, as may be the case for responses associated with climate change, for example. Thus research for monitoring and modeling approaches that can accurately predict ecosystem responses are highly merited.

Charge Question 3: Addressing *Cladophora* growth.

As noted in the reports, the association between bioavailable P in the water column and *Cladophora* growth and subsequent shoreline fouling are far leas clear than the associations between P and phytoplankton, cyanobacteria and hypoxia. Thus no P loading targets for this indicator are proposed at this time. Rather, more research to identify the best input variables (such as P contents in *Cladophora*) is emphasized, which appears reasonable. On a positive note, it is appears to be concluded in the Annex 4 report that reductions in whole lake P loading to achieve the hypoxia threshold may achieve the *Cladophora* goal. Thus it will be important to determine *Cladophora* responses to reduced P loading in conjunction with hypoxia responses.

Charge Question 4: Consideration of nitrogen control.

It is acknowledged that N may also play a role in algal blooms and hypoxia in Lake Erie. However, a compelling justification is made for the emphasis on P, at least at this point. This includes the reasonable assumption that controlling P loading, especially from non-point sources, will also control N loading. Thus it seems a reasonable starting point to monitor indicator responses to reduced P loading, along with quantifying N loading. If the indicator responses meet the threshold objectives, then the assumption would appear to be supported. If not, the observed N data will become very important, and perhaps modeling efforts for N that parallels these for P will be a next step for nutrient management in the Lake Erie watershed.

Charge Question 5: Accounting for inter-annual variability in hydrology in assessing reduction of tributary loadings.

I leave this to the hydrologists in the room!

Charge Question 6: Evaluation of nutrient reduction goals for Lake Erie.

This question was largely addressed above, with Charge Question 2. The P loading targets for each basin appear reasonable and best estimates available given existing data and models. As acknowledged in the reports, an adaptive management approach, informed by sophisticated monitoring, research and model development, is essential for the long term protection of the Lake Erie ecosystem. An interval for reassessments every 5 years seems reasonable, particularly given lag times between implementation of P management strategies and impacts on P loading and concentrations, and indicator responses, as well as funding realities. However, ideally a part of the adaptive management strategy would be for shorter intervals if, for example, P loading is not responding as desired by management practices, or environmental variables such as climate change are accelerating or altering responses in ways unanticipated by the current ensemble modeling approach.

Mr. Douglas Endicott

Preliminary comments in response to the charge questions for the upcoming Lake Erie Phosphorus Objectives Review Panel meeting

The EPA requests Science Advisory Board (SAB) review of the current modeling results and other information used to inform development of the binational phosphorus reduction targets. We are seeking a critical review so that we can ensure the Agency's ongoing efforts to develop, implement and evaluate nutrient reduction goals for Lake Erie are based on sound scientific data, analyses, and interpretations. In a spirit of adaptive management, we are most interested in SAB advice on enhancements to the modeling approach, or new approaches to consider, that will help us proactively manage eutrophication issues in Lake Erie in the long term.

Charge Questions:

Approach for Developing Lake Erie Phosphorus Load Reduction Targets

Nine different Lake Erie models were used to predict the response of selected eutrophication response indicators to different phosphorus load scenarios (see Table 1 in the Annex 4 Ensemble Modeling Report). The eutrophication response indicators evaluated were (1) overall phytoplankton biomass represented by chlorophyll a, (2) cyanobacteria blooms in the Western Basin, (3) hypoxia in the hypolimnion of the Central Basin, and (4) *Cladophora* in the nearshore areas of the Eastern Basin. Technical evaluation criteria were used to assess the capabilities of each model (see Section 2.3 and Appendix B of the Annex 4 Ensemble Modeling Report) and load-response curves were generated for each eutrophication response indicator (see Section 3 and Appendix B of the Annex 4 Ensemble Modeling Report).

1. *Please comment on whether the evaluation of the models was adequate to inform how model results should be interpreted, given differences in model complexity and scale. Please identify any additional analyses that may be needed to improve future development and interpretation of the load-response curves for the eutrophication response indicators.*

The Ensemble Modeling Report and Recommended Phosphorus Loading Targets document clearly present the following significant findings: (1) Cyanobacteria biomass reduction requires a focus on reducing TP loading from the Maumee River, with an emphasis on March-July high-flow event loads; (2) reducing central basin hypoxia requires annual load reductions to the central+western basin greater than those needed to achieve cyanobacteria biomass reductions; and (3) loading reductions targeting cyanobacteria blooms and hypoxia will improve eastern basin Cladophora densities.

A good effort was made to evaluate the models using several criteria, however this was obviously challenging given the significant differences between the models in terms of structure, resolution, and processes used in their calibration, confirmation and application. The modelers relied upon existing models, built and tested with a range of conditions. The models were not compared using identical inputs and measured response data sets, although there was commonality between sampling programs. Given these limitations, the load-response curves shown appeared to be as comparable as possible. Standardized model performance metrics were presented, although visual fit and statistical comparisons used to judge performance tended to be qualitative.

Discussion of the model evaluation criteria indicated variable extent/quality of calibration and confirmation between models. There were also considerable differences in the degree to which parameter sensitivity and predictive uncertainty were addressed.
Comments about specific models

Total Phosphorus Mass Balance Model: Expansion of model to predict chlorophyll and central basin hypoxia via empirical relationships with TP was mentioned, but it did not appear to be used.

EcoLE: Figure B5-7 suggests lack of fit for chlorophyll concentrations in the western basin.

9-Box Model: Although an empirical chlorophyll-TP relationship was mentioned, it did not appear to be used. It was mentioned that spatial resolution of this model limits the ability to assess changes in long-term hypoxia. This model extrapolated hypoxic area using the Zhou formula based on summer average bottom DO. It is not clear what the error bars in Figure B6-8 (and subsequent figures) represent.

WLEEM: There is a very useful discussion of how model performance and calibration targets change depending upon the state variable: model accuracy is expected to be less for water quality (nutrient) parameters relative to physical parameters (temperature), and biological parameters (chlorophyll and Microcystis) will be less accurate than water quality. It appears that DO predictions also fall into this “less accurate” category of state variables. This analysis is probably applicable to all of the mechanistic models.

ELCOM-CAEDYM: Documentation suggests that the evaluation of this model was not completed within the timeframe of this project. No confirmation presented for DO.

Great Lakes Cladophora Model: The Cladophora model was applied generically to conditions in the eastern lake basin and used to develop a load-response curve; however, it was not used to set a loading target for TP. Performance metrics, sensitivity and uncertainty were not presented, apparently because it has not been calibrated or confirmed to data for Lake Erie. Because this is a process model, several steps and other models are required to establish quantitative relationship between TP loads and a Cladophora indicator. A TP load to Lake Erie of 7000 MT was suggested to result in

offshore DRP concentration that should not stimulate nuisance growth. As offshore DRP is reduced, control of cladophora growth would shift to direct nearshore inputs. Coupling the Cladophora model to a model capable of resolving nearshore P gradients such as ELCOM-CAEDYM was suggested, however this was apparently not done.
The full effect of nutrient load changes on DO could not be seen with the short (i.e., annual) simulations made with most of the models, so SOD rates were adjusted to capture nutrient load reductions. Further evaluation and comparison of how the SOD fluxes were adjusted for different loading scenarios in the models predicting hypoxia should be considered, given the importance of this process to the model predictions.

The document, R*ecommended Phosphorus Loading Targets for Lake Erie* describes the process followed by the Annex 4 Objectives and Targets Task Team to develop phosphorus loading targets for Lake Erie. The document indicates that, to achieve a Western Basin cyanobacteria bloom biomass threshold no greater than that observed in 2004 or 2012, 90% of the time, a spring Maumee River load of 860 metric tons of total phosphorus and 186 metric tons of dissolved reactive phosphorus is recommended. In addition, a 40% reduction in the spring load of total phosphorus and dissolved reactive phosphorus from other Western Basin tributaries and the Thames River is recommended. To meet a threshold of 2.0 mg/L or higher of hypolimnetic dissolved oxygen, an annual total phosphorus load of 6,000 metric tons to the Western and Central Basins is recommended. The Task Team did not recommend new phosphorus concentration objectives for the open waters or the nearshore be identified at this time.

2. *Please comment on whether the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives defined in the GLWQA (as reflected in Table 1 on page 7 of the document titled Recommended Phosphorus Loading Targets for Lake Erie).*

The selection of P loading targets from the model load-response curves was opaque and somewhat confusing. For example, Table 10 in the Conclusion section of the Ensemble Modeling Report presents TP loads predicted by each model to be associated with example ERI thresholds that are different from the thresholds used in the Recommended Phosphorus Loading Targets document. Different TP loads were predicted by each model for all of the ERIs; how was one target loading selected for each? The process of selecting P loading targets needs to be transparent and clearly understandable. Hopefully this will be presented at the June meeting, and can be addressed in future revision of these documents.

The recommended loading targets appear to reflect a consensus of the models and are the "best available estimates" of the load reductions needed to achieve the ERIs. The ensemble modeling process incorporated measures to harmonize some of the discrepancies between the models and improve agreement in the load-response curves. Uncertainty of model predictions, which were explicitly calculated for several of the models, was disregarded in selecting the loading targets. In some cases, these

uncertainties are quite large, especially for biotic ERIs (chlorophyll and cyanobacteria) and DO. The ensemble approach may help to constrain predictive uncertainty, however it is still there. The Task Team acknowledges this, and proposed to address uncertainty through an adaptive management process.

Practically, it may be unreasonable to base loading targets on a 95% confidence limit, when even the expected predictions indicate that very significant nonpoint nutrient loading reductions are called for to achieve ERI objectives. In other words, we may be forced to target TP load reductions without a "margin of safety". Still, it would be unwise to ignore instances where specific model predictions appear to be outliers in comparison to the ensemble consensus, and suggest that greater loading reductions may be necessary to achieve ERIs. Three examples from the Ensemble Modeling report stand out: (1) the U-M/GLERL HAB model's increasing susceptibility of western basin to cyanobacteria blooms over time; (2) the 9-box model's predictions of greater hypoxic area than predicted by the other models; and (3) the GLCM Cladophora biomass likely depending on nearshore P availability and cycling, not just eastern basin spring DRP.

Cladophora Growth

Additional phosphorus load reductions may be necessary to reduce nuisance levels of *Cladophora* in the nearshore waters of the Eastern Basin of Lake Erie. The Annex 4 Objectives and Targets Task team did not recommend a specific phosphorus objective or loading target to address *Cladophor*a growth. EPA and Environment and Climate Change Canada convened a workshop in January 2016 to assess the current state of knowledge of *Cladophora* growth in the Great Lakes and identify potential options for nutrient target development to be considered by the Annex 4 subcommittee. (Please see the background document titled "State of the Knowledge of *Cladophora* in the Great Lakes. Executive summary of Workshop held at NOAA-Great Lakes Environmental Research laboratory, January 26-28, 2016.")

3. *Please comment on whether scientifically-sound phosphorus load reduction recommendations to address Cladophora growth in the Eastern Basin of Lake Erie could be developed at this time.*

A phosphorus load reduction to address Cladophora growth in the eastern basin *was* presented in Appendix B-9 of the Ensemble Modeling Report. Given the scientific uncertainty as discussed in the background document, and the lack of a tool to predict nearshore nutrient concentrations, this load reduction may be more uncertain than those recommended for the other ERIs. If the 7,000 MT TP loading cannot be recommended because it does not address localized nutrient sources to the nearshore, then maybe the 0.9 μg/L SRP or 6.3 μg/L TP concentrations should be considered as nearshore phosphorus concentration objectives, at least until such time as a linked GLCM-nearshore nutrient model can be developed and tested. It is interesting to note that the 6.3 μg/L TP concentration is quite close to the 6 μg/L expected in the open waters of the eastern basin

with achievement of the target loadings, suggesting there may be little margin for nearshore nutrient gradients without stimulating Cladophora growth.

Nitrogen Control

While the current strategy focuses on limiting phosphorus loading to the Lake (total and dissolved forms) as the key mechanism for controlling excessive algal growth, it is implied or assumed that nitrogen loading likely will also be reduced through implementation of agricultural best management practices, and the Task Team recommended that tributary nitrogen loads to the Lake be tracked in addition to phosphorus.

4. *What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?*

Evaluation of Nutrient Reduction Targets

The inter-annual loading trends for the Maumee River are greatly influenced by annual variability in flows. The Objectives and Targets Task Team identified a maximum flow below which the target load should be met and recommended the use of flow-weighted mean concentrations (FWMC) as a benchmark for any given tributary target load.

5. *Please comment on the use of FWMC and any other approaches that should be considered to account for inter-annual variability in hydrology in assessing progress in reducing tributary loadings of phosphorus to the Lake.*

Flow Weighted Mean Concentrations (FWMC) are suggested as a useful means to address interannual variability by normalizing the tributary phosphorus loading/delivery with respect to flow, so that year-to-year performance (I assume this is referring to nonpoint source nutrient controls) is not confounded by interannual hydrology. The Objectives and Targets Task Team recommends using tributary FWMC as a benchmark to track progress in load reduction. FWMC can be calculated for tributaries by dividing the phosphorus load during a specified period (e.g., annually or March-July) by the cumulative flow during that period. FWMC calculation can be made for all western basin tributaries, assuming phosphorus loads and flows have been monitored.

FWMC was demonstrated by Stow and Borsuk (2003) as useful to aid in assessment of TMDL implementation on the Neuse River. Examination of Figures 10 and 11 in the Recommended Phosphorus Loading Targets document (Maumee River flow-weighted mean TP and DRP concentrations and loads) suggests that the same trends are evident in FWMC and loading, especially for 5-year running averages. FWMC has intuitive appeal because it is a concentration, which may be easier to understand and communicate than

"mass loading". However, it cannot be calculated until average loading and discharge flow rate are known. If FWMC trends downward over time, that could be interpreted as a sign of progress in nutrient control, but what if the trend in discharge is increasing over the same time period? Can FWMC make tributary monitoring data look biased or "too good" in high flow years? Would focusing on FWMC put off acknowledging the need to possibly adjust loading targets to address changes in precipitation and discharge? Both pros and cons of the FWMC benchmark should be considered.
It was noted that annual discharge from the Maumee River is highly variable due to variations in the intensity, amount, and timing of precipitation. This variability is also an important factor leading to yearly differences in phosphorus loads. Similarly, discharge from spring to early summer (March-July) varies annually, and inter-annual variability during this period has been associated with variations in the size of the summer cyanobacteria bloom (Stumpf *et al* 2013, Obenour *et al* 2014), therefore tributary loadings during this "critical period" merit particular attention.
In 2008, the Maumee River 2008 FWMC was 358 μg/L P for TP = 358 μg/L P and 78 μg/L P for DRP. If target loads were achieved with proportional decreases to TP and DRP, the target FWMC would be TP = 215 μg/L P and DRP = 47 μg/L P.

Over the past 20 years external TP loads to Lake Erie showed large year-to-year variation but no clear long-term trend. Interannual variability is largely driven by hydro-meteorological conditions, which modulate the timing and magnitude of surface runoff and ultimately the amount of nutrients delivered to the lake by tributaries. The inter-annual loading trends for the Maumee River are greatly influenced by annual variability in flows. Maumee River discharge increased from 1984-2013, a pattern that has been shown to be consistent with long-term precipitation increases (Stow *et al*. 2015). The Objectives and Targets Task Team has attempted to account for this confounding behavior in how it has identified a maximum flow below which the target load should be met and by recommending the use of FWMC's to track progress for any given tributary target load.

I think this quote from Stow et al. (2015) is particularly relevant to the discussion of the use of FWMC or other approaches that should be considered to account for inter-annual variability in hydrology:

While it is generally acknowledged that targets may be exceeded during years of unusually high precipitation and tributary discharge, the use of load targets remains a common management tool. However, Milly et al. (2008) highlighted the growing recognition that, for variables such as tributary discharge, the assumption of stationarity, in an era of uncertain climate change, poses management challenges. Our results, indicating progressive precipitation and discharge increases in the Maumee River basin and concurrent phosphorus input increases to Lake Erie, suggest that imposing fixed load targets may require phosphorus concentrations to be persistently lowered to compensate for increasing discharge, if the targets are to be achieved. As phosphorus load targets are re-evaluated pursuant to the updated 2012 GLWQA, it may be appropriate to address the

possibility that continued discharge increases into the future may affect target attainment even if phosphorus reduction strategies are successful.

The Task Team recommended development of a comprehensive adaptive management program that would include annual routine monitoring of appropriate load, FWMC, and in-lake nutrient-eutrophication response indicators in conjunction with an intensive monitoring, research, and operational model application program every five years.

6. *Please comment on the value of applying the existing eutrophication models on an ongoing basis to periodically evaluate phosphorus loading targets and eutrophication response indicators. What key elements should be included in the adaptive management approach to successfully implement and evaluate our nutrient reduction goals for Lake Erie?*

For an adaptive management program to succeed, it will be necessary to apply and update the models regularly (annually or biannually). This is necessary, as model refinement based on changing loadings and other forcing functions, and the availability of new data for model testing, are components of any successful model-based management program. In part, this is a recognition of the significant inherent uncertainties in the models. The adaptive management process should track the response of the system to changing loadings and forcing functions, evaluate the effectiveness of management efforts, and update management recommendations as more is learned about the processes underlying the system response. Adaptive management must include a monitoring program capable of tracking precipitation, streamflow and loading trends over time and in-lake responses, as well as studies directed at learning more about processes that may be important, but are incompletely understood.

The monitoring program needed for the adaptive management approach must be capable of detecting critically important changes in precipitation patterns and phosphorus loads. If the frequency of severe storms increases in the future as predicted by climate change models, phosphorus loading to the lake will increase and the frequency and severity of cyanobacterial blooms will also increase, requiring larger phosphorus load reductions than those in the Recommended Phosphorus Loading Targets document. If appropriately designed, funded and executed, the adaptive management approach could detect these occurrences and allow for modification of target loads. However (according to the Task Team recommendations), a five-year interval on focused science and monitoring will not be sufficient for effective adaptive management.

According to Dolan and Chapra (2012), the standard error of TP loading estimates to Lake Erie are already quite small (3.5% of the load), although between 1994 and 2008, an average of 35% of the tributary loads were from unmonitored tributaries. Adaptive management should look for opportunities to initiate monitoring on additional tributaries, while at the same time preventing the loss of current tributary monitoring capabilities. For example, the major tributaries to Lake Erie, including the Maumee, Sandusky and Cuyahoga Rivers, are often sampled more than once a day. Dolan and Chapra (2012) also

note the need for better measurements of dry deposition fluxes for phosphorus. Other areas for improvement would be to address possible inaccuracy in the nutrient loadings from the Detroit WWTP, and possibly speed up MDEQ's 5-year lag in reporting of data from their monitoring of the Detroit River.

In addition to annual hydrology variability, variability in wind stresses and the heat balance are also factors that contribute to the lake's response. In particular, the models predicting hypoxia emphasize the importance of the physical factors (temperature, wind, currents) that have a significant impact on the stratification timing and magnitude and the volume of the associated hypolimnion. This creates potentially significant variability in hypoxia as a function of a given phosphorus loading profile. The effect of variable physical factors was demonstrated by the application of the 1-D hypoxia model to the observed physical conditions for the years of 1987 – 2005. The results demonstrate approximately a factor of two variability in hypoxic area and mean hypolimnion DO concentration for a given TP load due to variation in physical conditions. Phytoplankton growth may also be expected to increase due to warming lake temperatures and reduced extent and duration of ice cover.

As noted in the Recommended Phosphorus Loading Targets document, there is increasing evidence that climate change in the Great Lakes will cause an exacerbation of the ERI's due to changes in all these factors. Monitoring these climate change impacts on the load-response relationships will be an important and necessary aspect of a good adaptive management program.

Mr. James Fitzpatrick

Lake Erie Phosphorus Objectives
Response to Charge Questions

General Comment

The Annex 4 Modeling Subgroup developed a consistent and logical approach towards developing phosphorus loading targets for Lake Erie. The draft Annex 4 Ensemble Modeling Report and the draft paper by Scavia, DePinto and Bertani provides a good overview of the approach and its limitations used to develop the objectives and target loads, as well as an overview of each of the models. It is interesting that despite the differences between and limitations of the models, the loading targets or response curves estimated by the models are quite similar for the various Environmental Response Indicators (ERIs). In performing my review and in developing my response to the charge questions, I attempted to identify model formulations or model coefficients which might call into question the validity or utility of the model in providing outputs (i.e., response curves) for the ensemble modeling approach. This is not meant totally invalidate the analysis, but to support the recommendations contained in the Recommended Phosphorus Loading Targets report and in Scavia, DePinto, and Bertani concerning additional field and process monitoring, continued updating and refinement of the models, and updating management recommendations using an adaptive management approach.

Charge Question #1

Given apparent time and budget constraints, which perhaps limited a full update and recalibration of the more mechanistic models (1-D Central Basin Hypoxia Model, ELCOM-CAEDYM, EcoLE and the 9-Box Model), I believe that the evaluation of the models was adequate to inform the application of the models to address the Western Basin cyanobacteria biomass ERI. I am less confident in the models as applied to addressing the Central Basin Hypoxia issue. The reason for this concerns the model formations and coefficients used to determine SOD, potentially a major sink of dissolved oxygen, together with the bacterial oxidation of particulate and dissolved organic carbon in the hypolimnion, during stratified conditions. In particular, the 1-D Central Basin Hypoxia Model, EcoLE and ELCOM-CAEDYM all utilize non-linear responses to P load reductions. Perhaps, this may be appropriate if one considers that at very high P loading rates, phytoplankton growth becomes light limited and the amount of algal carbon generated by photosynthesis, which can in turn settle to the bottom sediments and contribute to SOD, is capped. In the case of the 1-D Central Basin and EcoLE models, the hyperbolic function used to relate SOD to P load, results in SOD going to zero as the P load goes to zero. However, in the ELCOM-CAEDYM model, the SOD is still non-zero (and non-trivial) even when the P load goes to zero. It is not clear from the model descriptions whether or not atmospheric loads are accounted for separately in any of the models, so it is difficult to assess whether the SOD should go to zero at zero P load as is the case for the 1-D Central basin and EcoLE models (no, if atmospheric loads are still in the model and not removed) versus ELCOM-CAEDYM. Perhaps this can be discussed with the ANNEX 4 modeling team at next week's meeting before a final conclusion concerning the central basin hypoxia target loads can be finalized.

Charge Question #2

The recommended targets for western basin cyanobacteria appear reasonable and do appear to reflect the best available information. Again the consistency in the loading response curves seems to provide the most confidence in supporting these targets. There is less certainty with respect to the central basin hypoxia question. Again, this hinges to some degree on the way SOD is handled in the models, which in turn may be reflective of the fact that several of these models are dated with respect to the current state-of-the-science with respect to eutrophication models. It also is influenced by the fact that the proposed nutrient targets to help reduce hypoxia differ in magnitude by almost a factor of two (versus ~30% for the cyanobacteria targets). The current state-of-the-science in eutrophication modeling is to include explicit sub-models of sediment nutrient diagenesis and SOD and nutrient flux sediment nutrient flux model – SFM). Perhaps an adaptive management approach, which starts with a lower P target than the proposed average P load target would be appropriate for consideration as (1) additional field and monitoring of SOD and sediment P flux are conducted and (2) updating the central basin eutrophication models to include a SFM and re-calibration of the models should be considered.

Charge Question #3

Given the current state of the models with respect to spatial resolution in the eastern basin and the current state of the science with respect to modeling Cladophora, I support the modelers and Task Team's recommendation not to support firm targets for this ERI at this time.

Charge Question #4

As noted it is likely that agricultural BMPs will likely reduce nitrogen as well as phosphorus. While my expertise is not in agricultural practices or agricultural BMPs, one might ask whether the BMPs will preferentially reduce N over P. If so, there may be a potential, perhaps very, very small, that the ratio of P to N may change dramatically. If so, there is a chance, again likely small, that N could become co-limiting with P. While this may an overall benefit, it could also result in some unintended changes in community composition. To address this concern, a literature review could be conducted which would gather information as to expected reductions in P and N from implementation of various agricultural BMPs. If the literature suggests that there is not a preference for N removal over P, then it is unlikely that additional analysis/monitoring needs to be conducted. If that is not the case, then additional monitoring of inorganic nitrogen and perhaps nutrient addition (P-alone, N-alone, P and N) experiments on lake algae may be considered.

Charge Question #5

The use of flow-weighted mean concentrations, together with flow measurements, is an appropriate metric for assessing progress in reducing tributary P loadings to the lake. However, I don't know if some storm event composite sampling of TP and SRP has been conducted in the tributaries. During large storm events it is possible that the concentrations of TP might be

influenced by resuspension of sorbed P from the river beds or possibly shore-line erosion. If so, the highest concentrations are likely to be found on the rising limb of a storm hydrograph. Therefore, if data do not exist to help evaluate this concern, perhaps some sampling or monitoring over the course of a storm or high flow event might be considered.

Charge Question #6

This panel member strongly supports the Task Team's recommendation of an adaptive management approach that includes applying eutrophication models on an ongoing basis. However, I would recommend that the recommendation be extended to include an update to some of the mechanistic models to include a SFM submodel rather than just use the existing eutrophication models. While this will require funding to support additional field and laboratory studies, as well as the cost to upgrade and re-calibrate the models, the required monies for these tasks are likely to be much, much smaller than the costs to implement nutrient management programs required to meet the proposed target P loads.

Additional Modeling Question/Comment

In reviewing the table of model coefficients used for the 1-D Central Basin Hypoxia Model I see that the P:C ratio is 0.01, which I assume is for phytoplankton. This corresponds to a C:P ratio of 100, which is well above the conventional Redfield ratio of 40. Perhaps, this is a conservative approach, since phytoplankton can approach the C:P ratio of 100 when they become nutrient limited. Comments from the Modeling Team concerning this choice of coefficient would be appreciated at next week's meeting.

Dr. Robert Heath

Question 1. *Was the evaluation of the models adequate to inform how model results should be interpreted?*

Given the wide range of model types, I believe it is appropriate to measure each of the models using the same statistical approaches, for the sake of uniformity in evaluation. I recognize that some types of models are amenable to a rigorous uncertainty analysis that others are not. The statistical performance measures chosen (root mean square error, percent bias, and mean absolute relative error) are commonly accepted measures to assess model performance when compared against observations. I also agree with the presentation of the model outputs as a result within a 90 or 95 percent confidence envelope.

Question 2. *Do the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives as reflected in Table 1 of Recommended Phosphorus Loading Targets for Lake Erie?*

It is rational to base important decisions, such as the recommended target loads to Lake Erie, on a wealth of information gathered over a span of 40 - 50 years. Many of the models in the chosen ensemble were developed in the 1980s based on existing information regarding nutrient concentrations and processes known at that time. But basing these decisions on long-standing observations carries the implicit assumption that things are now as they were then. Specifically, to manage a system by controlling external P-loading assumes that external P-loading is the major driver of the re-eutrophication of Lake Erie since 1995. That assumption was correct in 1978 when GLWQA was adopted, but there is an increasing wealth of data indicating that that assumption is incorrect. Even the fact that P-loadings have been lower than the mandated 11,000 MTA in all but three years since 1995 during the re-eutrophication process argues that the ecosystem has changed and to manage it requires a thorough understanding of its structure and function now.

I believe there are significant reasons to look beyond the notion that Western Basin cyanobacterial blooms are driven primarily by external P-loading and that hypoxic events in the Central Basin are driven in turn by those cyanobacterial blooms. Briefly, here are three items that require further investigation:

a. With the advent of dreissenid mussels it may well be that the significance of **internal loading** as a source of P has increased in relative importance not reflected in the current models. Also, the models are based on observed load-response curves and may well represent responses to events that are coincident with the observed external P-loading.

For example, much of the external P-loading results from runoff following storm events; yet those same storm events may well exacerbate turbulent and advective loading from the sediments. Or the external P-loading may also be accompanied by external N-loading from agricultural fertilizers. Ongoing research indicates that internal P-loading may be more important than generally considered as conventional wisdom.

b. The **role of nitrogen** may well be more important than generally regarded. The models are predicated on the explicit or implicit assumption that phytoplankton are strongly P-limited and will respond directly to available phosphorus (e.g. SRP aka DRP). This assumption disregards the well- known fact that many phytoplankton physiologically respond more to N:P than to either N or P separately. Numerous studies have shown that the combination of N and P often results in higher cyanobacterial biomass than either nutrient added singularly (e.g., Elser et al. 2007; Lewis & Wurtsbaugh 2008; Scott & McCarthy 2010 & 2011). With increasing frequency since 2002 there have been reports of the algal blooms being N-and-P co-limited or N-limited, especially during mid- to late summer.

c. The role of the **winter-spring diatom bloom** in causing regions of hypoxia in the Central Basin of Lake Erie has been generally neglected. The early spring diatom bloom is not P-limited, rather it's limited by light and Si. Recent work by Reavie et al. indicates that Central Basin hypoxia is driven by the winter-spring diatom bloom, which is P-replete (N:P = 3.1) and light- and Si-limited in growth. Diatom biomass was 1.5 to 6 times greater than summer algal bloom biomass and was able to account for oxygen demand in the profundal waters. The link between cyanobacterial blooms in the Western Basin and hypoxia in the Central Basin has been indirect (at best) and largely anecdotal.

Question 3. *Can scientifically-sound recommendations for external P-load reduction be made at this time to address Cladophora growth along the shores of the Eastern Basin of Lake Erie?*

Great Lakes *Cladophora* Model (GLCM) The GLCM is a mechanistic, mass balance model with two state variables, net algal biomass (growth minus respiration and sloughing) and stored P. The forcing conditions are available SRP, incident light intensity and temperature. The model was calibrated by direct observation in field (Lake Huron) and laboratory studies; it was confirmed by comparing observations in Lake Michigan with fit to model predictions. Sensitivity analysis showed that the model was most sensitive to the minimum cell P-quota, the maximum growth rate and the maximum respiration rate; it was marginally sensitive to parameters related to phosphate uptake. From model curves for (SRP vs. maximum standing crop) and (SRP vs. stored P content) it was shown that SRP of 0.9 µg P/L would yield a maximum standing crop of 30 g DW/m^2 and a stored P content of 0.075 percent P. That is, at 0.9 µg P as SRP/L would yield an acceptably low standing crop and low growth potential for *Cladophora*. This level of SRP was related to TP concentrations and total P load to Lake Erie via load-response curves derived empirically (Fig B9-2 and B9-3 in Appendix B-9) to imply

that the *Cladophora* growth and P-quota could be met with a total TP load reduction to 7,000 MT/yr, or a load reduction of 25 percent. Because the goal of 40 percent reduction in total TP load to Lake Erie appears to be necessary to attain other desired ERIs, the implication of the GLCM is that by meeting those goals, the goal of adequate reduction of *Cladophora* growth will be met also, *de facto*.

Caveats On the face of it, the GLCM appears to provide adequate answers sought regarding attainment of the ERI of reduction of *Cladophora* standing crops to acceptable levels and with little growth potential, as indicated by a low P-quota. Yet, I believe that further considerations need to be made before we can answer Question 3 affirmatively.

- The GLCM was calibrated and confirmed on Lake Huron-Michigan. I believe that it should be confirmed in the Eastern Basin of Lake Erie because there are significant differences between Lake Erie and the lakes on which this model was developed.

- *Cladophora* growth depends on SRP content *in situ*, which in turn depends not only on the TP levels in the open lake (as this model considers) but also on local inputs from nearby tributaries, as well as the presence of dreissenids. I believe that a scientifically-sound model necessitates site-specific considerations, including local hydrodynamics and meteorology.

- Although this model considers *Cladophora* as the only issue of concern, it needs to recognize that there are other nuisance benthic algae (NBA, e.g. *Chara*, *Lyngbya*, *Spirogyra*, etc.) that can cause similar problems. Consideration of the similarities and differences among these various species needs to be made to provide an adequate representation of the problems of NBA in general.

- The nuisance attribute of *Cladophora* is actually the formation of "beach muck" and the attendant problems that arise from it; the formation of "beach muck" is initially from sloughing of standing crops of the benthic alga. The GLCM provides only a crude estimate of this process, modeling sloughing as a constant coefficient of the calculated standing crop. The process or processes that lead to sloughing (local hydrodynamics, algal senescence, etc.) and eventual decay to "beach muck" need further investigation and likely need to be appended to the GLCM.

Question 4. *What questions, relationships, or research priorities related to N-loading in its various forms and sources (e.g. nitrate-nitrite, ammonium, urea) and in-lake cycling must be addressed?*

With increasing frequency, reports have shown that N may be crucially involved in the processes that yield cyanobacterial blooms in Lake Erie. In contrast to the P-cycle, where only oxidation state +5 is important, in N-dynamics many oxidation sates need to be considered along with the

redox processes involved. The questions regarding the significance of N go beyond simply how much is loaded but also to processes in N-dynamics. My responses to this question are informed by my colleagues Silvia Newell and Mark McCarthy (Wright State Univ.). Several critical questions need to be addressed to understand the role of N in the Lake Erie ecosystem.

- To what extent can denitrification mitigate excess nutrient loading?
- How much nitrogen is removed vs. recycled by sediment microbial processes?
- How rapidly is reduced N (ammonium and urea) recycled vs lost to the nitrate pool through nitrification?
- What is the reduced N demand of both phytoplankton and bacteria and can water column turnover and/or sediment efflux meet that demand?
- What is the N budget for Lake Erie, especially the Western Basin? If all inputs were completely reduced, how long would legacy nutrients drive the bloom?

<u>Rationale for controlling N inputs to Lake Erie:</u>

The Maumee River drains a mostly agricultural watershed and discharges into the Western Basin of Lake Erie, where annual cyanobacterial blooms have occurred since the mid 1990s. These blooms are fundamentally different from those occurring in Lake Erie prior to P-loading reductions implemented in the 1970s. While most blooms prior to the 1990s were comprised of filamentous, heterocystous cyanobacteria (e.g., *Aphanizomenon*, a potential nitrogen N-fixer), the modern blooms are comprised mostly of the non-N-fixing genera *Microcystis* (reviewed by Steffen et al. 2014). The inability of these cyanobacteria to fix atmospheric N suggests an important role for external N loads from the watershed as well as an essential role of internal N recycling mechanisms in modulating the total biomass and especially the composition of the cyanobacteria community in the Western Basin. In addition, the toxin produced by *Microcystis* (and other cyanobacteria), microcystin, is comprised of a large proportion of N (10 N atoms per molecule), and production of microcystin is strongly correlated with available N (Davis et al. 2015). This apparent N problem in Lake Erie is not confined to the *Microcystis* blooms in the western basin. Indeed, algal issues in other parts of the lake, including annual *Planktothrix* blooms in Sandusky Bay (Davis et al. 2015) and ongoing blooms of *Cladophora* (Davies & Hecky 2005), also involve non-N-fixing algae.

Despite the absence of N-loading reduction recommendations in many documents urging additional control of external P-loading in the Lake Erie watershed (e.g., Stumpf et al. 2012; Michalak et al. 2013; IJC 2014; Scavia et al. 2014; Dove & Chapra 2015), there is mounting evidence that N control is also needed (Chaffin et al. 2013; Davis et al. 2015). Indeed, arguments for N control on a global scale are gaining significant traction, since the massive

increase in reactive N levels, primarily due to Haber-Bosch N fixation for fertilizer production, has been linked to loss of biodiversity, climate change, water and air pollution, and aquatic hypoxic areas (Erisman et al. 2015). Unfortunately, the reality is that most impaired ecosystems have already achieved the simulated conditions of simultaneous N and P additions. If only N or P were reduced, excess biomass levels would still exist. We can no longer afford to ask whether we have an N or a P problem—we have a nutrient problem and we need to address it as such.

Question 5. *Does the use of flow-weighted mean concentrations (FWMC) provide an adequate means to assess progress in reducing tributary P-loadings to Lake Erie?*

It is well known that external P-loading to Lake Erie is largely driven by storm events, such that P-load is generally greatest in wet years and least in dry years. The inter-annual external P-load varies widely as years vary between wet, average, and dry. Judging success of the recommended management plan would be difficult if only loads were considered (I.e. the signal-to-noise ratio would be too high to see changes readily from the noisy baseline of past annual P-loading). Scaling P-load for volume of water carrying that load (i.e. reporting the flux as an average concentration) is an acceptable way of trying to smooth the response curves and reduce the signal-to-noise ratio, as shown in Fig 10 – 13 (*Recommended Phosphorus Loading Targets*) . So, yes, I believe that using the FWMC is an acceptable variable to track as a means of assessing the success of the management plan to reduce external P-loadings.

Question 6. *What key elements should be included in the adaptive management approach to successfully implement and evaluate the nutrient reduction goals for Lake Erie?*

Evaluation of the nutrient reduction goals should be done in the most direct way possible. Evaluation of success should include reduction in FWMC of external P-load, as well as external P-load directly. Also, N-loading in its several important forms should be monitored: nitrate-nitrite, ammonium, urea, as well as investigation of the various redox processes need to be research to a greater extent, as cited above.

As indicators of success of the management plan, chlorophyll *a* should be monitored, as well as standing crop of cyanobacterial biomass in the Western Basin, and extent of hypoxia in the Central Basin. Because one of the aims of the management plan is to maintain the Western and Central Basins as mesotrophic and the Eastern Basin as oligotrophic, I suggest that the Trophic State Index [Carlson, R.E. 1977. Limnol. Oceanogr. 22(2) 361-369] be determined at selected stations in each basin in August. This TSI is based on the assumption that the phytoplankton are strongly P-limited. If they are, the TSI determined from chl *a* or Secchi depth or Total P will be similar values. If the phytoplankton are not P-limited the TSI_{chl} will differ from TSI_{TP} and TSI_{Secchi}. I suggest that research continue to work toward an adequate model for predicting *Cladophora* at selected sentinel sites, as suggested in the reports. I also believe that research needs to be done to corroborate the recent studies that suggest that the winter-spring diatom bloom may be responsible for development of hypoxic regions in the Central Basin.

Dr. J. Val Klump

Preliminary Responses (6-15-16) submitted by: J. Val Klump, University of Wisconsin-Milwaukee

Charge Questions:

Approach for Developing Lake Erie Phosphorus Load Reduction Targets

Nine different Lake Erie models were used to predict the response of selected eutrophication response indicators to different phosphorus load scenarios (see Table 1 in the Annex 4 Ensemble Modeling Report). The eutrophication response indicators evaluated were (1) overall phytoplankton biomass represented by chlorophyll a, (2) cyanobacteria blooms in the Western Basin, (3) hypoxia in the hypolimnion of the Central Basin, and (4) *Cladophora* in the nearshore areas of the Eastern Basin. Technical evaluation criteria were used to assess the capabilities of each model (see Section 2.3 and Appendix B of the Annex 4 Ensemble Modeling Report) and load-response curves were generated for each eutrophication response indicator (see Section 3 and Appendix B of the Annex 4 Ensemble Modeling Report).

1. *Please comment on whether the evaluation of the models was adequate to inform how model results should be interpreted, given differences in model complexity and scale. Please identify any additional analyses that may be needed to improve future development and interpretation of the load-response curves for the eutrophication response indicators.*

General comment: The evaluation of individual models detailed both the strengths and weaknesses of each and was both sufficiently critical and thorough. The inter-model evaluations were obviously aided by the examination of a common set of metrics allowing comparisons of model results and outputs to be evaluated vis a vis one another on relevant time frequencies and spatial scales, despite differences in model approach, complexity and scale. In short, the evaluation process was adequate, the comparisons reliable, the uncertainities clearly identified and quantified, and the interpretations of results measured, accurate and critical.

With respect to future development and interpretation the following warrant consideration and additional research, and in most instances were identified as future needs:

- The relationship between toxin production and cyanobacterial production. As noted, not all algae produce toxins, and not all toxin producing algae produce toxins all the time. The trigger for toxin production is not well understood. In addition, algae other than cyanobacteria, also produce toxins, although they are not dominant in this system. The Obenour model suggests an increasing sensitivity of the western basin cyanobacteria to

loading over time. This is an unknown but could be a significant ecological evolution in the system based upon a variety of factors (loading, species abundance, seed populations, climate/weather, etc).

- Variability in discharge and its timing is clearly recognized as important. In this respect the longer term influence of projected climate change is highly relevant. A number of these models are climate change scenario adaptable, and future inclusion of mid and late century simulations would be useful and instructive both for examining the potential impacts on the dynamics within the ecosystem and for identifying potential longer term management considerations.

- For the areal extent of hypoxia in the Central basin there is less quantitative agreement among the models, although the nature of loading-response functions are coherent with one another. There is, for example, a significant difference in the predictions of the areal extent of hypoxia between versions of the EcolE model that appear to depend, in part, upon the depth/thickness of the bottom layer modeled – with the thinner layer (1.0 meters) projecting a 2-fold greater "hypoxic zone" than the 1-3 meter model layer. This makes sense in a system which is both shallow and highly stratified, and also argues for the interannual and long term projections of variability in wind fields, thermal stratification and warming, and currents – all climatically driven in the long run – and also recognized as important in the report. The approach of using bottom water oxygen concentrations as one metric of choice, in this regard, makes some sense, and the model predictions are in closer agreement with respect to DO concentrations. Although Zhou et al (2013) show that average bottom water concentrations less than 4 mg/L will result in significant hypoxia, the recommendation is for a 2 mg/L objective. A 4 mg/L objective is deemed too stringent for the load reduction required (and relates to the following comment on fish production). A 2 mg/L objective has several significant benefits, including a reduction in the areal "footprint" of hypoxia and benefits to the benthos, especially on the peripheries of the "dead zone".

- Climate change projections also call for extended summertime conditions and warmer air temperature with potential increases in the stratified period on the order of 4-6 weeks. This obviously impacts hypoxia by extending the period during which water column and benthic respiration can deplete oxygen. A lengthened stratified period could out weigh many of the other drivers or attenuators of hypoxia, exacerbating the hypoxia in the central basin despite reductions in loading.

- There is concern expressed over the impacts of loading reductions on a "healthy fish community", and presumably fish production, i.e. fish vs phosphorus. These models do not address this issue directly, other than to project mean annual phosphorus concentrations that would keep Lake Erie in a "mesotrophic" state. This deserves further research. Presumably much of the current algal production in the lake is an ecological dead end and does not enter the food chain, but simply goes to sink and fuels sediment respiration, increased oxygen

consumption, and a depauperate benthos. Hence, some reduction in loading, if it tilts the ecosystem to a more energetically available food chain, could be beneficial. This is a complex, however, and may involve inkages not yet well understood, e.g. response of the dreissenid community, etc.

The document, R*ecommended Phosphorus Loading Targets for Lake Erie* describes the process followed by the Annex 4 Objectives and Targets Task Team to develop phosphorus loading targets for Lake Erie. The document indicates that, to achieve a Western Basin cyanobacteria bloom biomass threshold no greater than that observed in 2004 or 2012, 90% of the time, a spring Maumee River load of 860 metric tons of total phosphorus and 186 metric tons of dissolved reactive phosphorus is recommended. In addition, a 40% reduction in the spring load of total phosphorus and dissolved reactive phosphorus from other Western Basin tributaries and the Thames River is recommended. To meet a threshold of 2.0 mg/L or higher of hypolimnetic dissolved oxygen, an annual total phosphorus load of 6,000 metric tons to the Western and Central Basins is recommended. The Task Team did not recommend new phosphorus concentration objectives for the open waters or the nearshore be identified at this time.

2. *Please comment on whether the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives defined in the GLWQA (as reflected in Table 1 on page 7 of the document titled Recommended Phosphorus Loading Targets for Lake Erie).*

General comment: Yes, the recommended targets do reflect the best available information, and indeed, include a relatively extensive set of historical data, both *in situ* and remotely sensed.

One aspect of the hypoxia problem which probably deserves more attention is the internal oxygen consumption capacity that has presumably built up and will continue to build up over time as excess organic matter delivered to and accumulated within the depositional environments of the central basin. Some of the models do include a diagenetic component, but I believe it is a steady state condition. The time dependence of improvement in hypoxia may depend, in part, upon the excess oxygen demand generated by years of enhanced deposition. An estimate of the residence time for “metabolizable carbon” in these sediments is probably on the order of 10-50 years (a guess). Evidence in some Swiss lakes suggests recovery from hypoxia following loading reductions is delayed by this internal reserve of accumulated oxygen consumption capacity.

Future climate conditions also warrant inclusion in the hypoxia modeling – changes in the length and intensity of stratification, alteration of regional wind climatology, lake level fluctuations, etc. All have the potential to influence the extent and duration of hypoxia on a longer term basis. One of the major predictions of climate models is an increase in variability, e.g. while lake levels may or may not change from current levels, the range of expected fluctuation in lake levels is

projected to increase, i.e. planning and modeling for both higher than historical levels and lower than historical levels may be warranted.

Cladophora Growth

Additional phosphorus load reductions may be necessary to reduce nuisance levels of *Cladophora* in the nearshore waters of the Eastern Basin of Lake Erie. The Annex 4 Objectives and Targets Task team did not recommend a specific phosphorus objective or loading target to address *Cladophor*a growth. EPA and Environment and Climate Change Canada convened a workshop in January 2016 to assess the current state of knowledge of *Cladophora* growth in the Great Lakes and identify potential options for nutrient target development to be considered by the Annex 4 subcommittee. (Please see the background document titled "State of the Knowledge of *Cladophora* in the Great Lakes. Executive summary of Workshop held at NOAA-Great Lakes Environmental Research laboratory, January 26-28, 2016.")

3. *Please comment on whether scientifically-sound phosphorus load reduction recommendations to address Cladophora growth in the Eastern Basin of Lake Erie could be developed at this time.*

No. The report has accurately described the difficulties of providing a load reduction recommendations tied to Cladophora growth. Models are largely pinned to DRP. In order to estimate DRP, the relationship between DRP (not measured) and TP (measured) is assumed to be some value (~ 1:7), the variability of which not well known, leading to additional uncertainty in the model. Nutrient supply and recycling and the confounding affects of Dreissenid mussels complicate the predictions of Cladophora growth and distribution, as well. Local inputs are important, however, regions of the Great Lakes remote from local inputs have experienced major Cladophora outbreaks. There is good evidence that offshore waters supply significant fraction of the nutrients to the nearshore through convection and mixing processes. Hence boundary layer effects, onshore/offshore transport dynamics and local influences and geomorphology – along with the non-steady state condition of Dreissenid populations and their multiple impacts on light climate, substrate and nutrient availability, make tying a whole lake or basin nutrient reduction target to Cladophora abatement highly problematic.

Nitrogen Control

While the current strategy focuses on limiting phosphorus loading to the Lake (total and dissolved forms) as the key mechanism for controlling excessive algal growth, it is implied or assumed that nitrogen loading likely will also be reduced through implementation of agricultural best management practices, and the Task Team recommended that tributary nitrogen loads to the Lake be tracked in addition to phosphorus.

4. *What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?*

First of all, this is a hotly debated topic within the scientific community with strong proponents on both sides of the nitrogen control issue. In part, the argument for considering nitrogen abatement relates to the fact that microcystis does not fix nitrogen, and denitrification can result in the loss of fixed nitrogen leading to nitrogen deficiencies, particularly late in the evolution of a HAB. Phosphorus is more efficiently retained within the system and subject to regeneration – hence may, indeed, "build up" relative to nitrogen. On the other hand, numerous species of algae are capable of nitrogen fixation, and hence, have an "unlimited" source. Even advocates for nitrogen control, however, do not advocate for nitrogen-only control. As pointed out in the report, many, if not most, agricultural BMPs that target phosphorus runoff abatement will also affect nitrogen abatement, albeit to varying degrees. The role of nitrogen cycling in the western basin, however, appears to be an open question, and the recommendation for further research and more extensive monitoring would appear both wise and warranted. Adaptive management should look at the impact on both phosphorus and nitrogen and their stoichiometric relationship as phosphorus abatement is pursued. Since nitrogen is an essential nutrient for agriculture that is extensively applied, reducing losses of this nutrient has benefits on the front end as well.

Evaluation of Nutrient Reduction Targets

The inter-annual loading trends for the Maumee River are greatly influenced by annual variability in flows. The Objectives and Targets Task Team identified a maximum flow below which the target load should be met and recommended the use of flow-weighted mean concentrations (FWMC) as a benchmark for any given tributary target load.

5. *Please comment on the use of FWMC and any other approaches that should be considered to account for inter-annual variability in hydrology in assessing progress in reducing tributary loadings of phosphorus to the Lake.*

The maximum March-July flow from the Maumee River is estimated to be below 4×10^9 m^3 90% of the time. This is based upon data from 1990-2013. Only twice in that period did the flow exceed that level (2003 and 2011). Since loading is heavily dependent on flow, it is logical to set a target load for a reasonable upper boundary in flow. The obvious question, of course, is whether this historical range is likely to hold in the future. This region is projected to be both warmer and wetter, with precipitation increases on the order to 5-15%. Clearly it is well recognized that it is not just the total flow that is important, but also the timing and intensity. Projections are for warmer, wetter winters and an increase in the frequency of intense precipitation events. These intense events tend to dominate loading with maybe 70% or more of

the total annual loads coming during a relatively small number (~10 or so) events. This argues for the inclusion of climate change scenarios. Increases in precipitation may not result in increased runoff in all circumstances, however, since warmer conditions can drive up evapotranspiration rates and offset flow.

Targets for tributary loadings of phosphorus should be based on the total flux of phosphorus to the lake, equivalent to concentration x flow. The use of FWMC is logical given sufficient monitoring data, and assumes flow is measured which is critical.

The Task Team recommended development of a comprehensive adaptive management program that would include annual routine monitoring of appropriate load, FWMC, and in-lake nutrient-eutrophication response indicators in conjunction with an intensive monitoring, research, and operational model application program every five years.

6. *Please comment on the value of applying the existing eutrophication models on an ongoing basis to periodically evaluate phosphorus loading targets and eutrophication response indicators. What key elements should be included in the adaptive management approach to successfully implement and evaluate our nutrient reduction goals for Lake Erie?*

To employ an adaptive management approach, it is essential to 1) collect sufficient data that can inform these models, and 2) refine and reapply the models based on that information, new data and new process information. These models represent a significant investment in understanding the response of this system. They will only get better, and we will learn as much when they fail as when they are successful in predicting ERI's.

In addition to the 17 recommendations given in the report (p. 45-46), a significant effort obviously needs to be made on the landscape side of this equation. Several of these models are capable of being and are being linked to watershed runoff models tied to land use practices on a relatively fine sub-basin scale. This is the upstream source function and any successful restoration plans will require a major effort in this regard – not only in data collection, monitoring and model development, but also in technology and best management practice development, stakeholder engagement, and regulatory, management, and policy agency participation combined with willingness to take needed actions, etc.

Dr. Douglas McLaughlin

Preliminary comments following initial review of:

Annex 4 Ensemble Modeling Report ("modeling report")

Recommended Phosphorus Loading Targets for Lake Erie ("recommendations report")

D. McLaughlin, 6/15/2016

My preliminary comments primarily respond to charge questions #1 and #6. The two documents provided to the SAB panel for review contain many elements that are responsive to advice previously given by this panel about evaluating and expressing uncertainty in load reduction estimates. For example, Section 3 of the modeling report provides information on statistical assessments of how well model predictions fit field data, and typically provides an explanation when certain statistical measures were not calculated. In addition, figures showing the load-response curves generated using different models usually include an element indicating prediction uncertainty and the selected numeric response threshold. Examples include Figures 11-13 in the modeling report (reproduced as Figures 14-16 in the recommendations report) which show the load response curves from three different models for cyanobacteria bloom size in the Western Lake Erie basin as a function of spring Maumee River total phosphorus load. Mean prediction lines and 95% prediction intervals are provided. Also shown is the selected Eutrophication Response Indicator (ERI) threshold for cyanobacteria of 9600 MT dry weight biomass (or equivalent as described in the report) used to distinguish between "mild" and "severe blooms". Together, these elements show where the load-response lines intersect with the ERI threshold, providing a useful summary of the central tendency of and the uncertainty in the estimated loads necessary to achieve the ERI threshold.

Figures such as these provide important information for evaluating and communicating the usefulness of these models for the specific decision-making objectives within the context of Annex 4, i.e., to identify the phosphorus load reductions necessary to achieve Lake Ecosystem Objectives for Lake Erie. They help translate information on statistical performance into a corresponding range of phosphorus loading estimates required to achieve a numeric ERI threshold, define regions of both stronger and more limited predictive capability, support more realistic expectations of the responses that are likely to be observed as specific phosphorus load reductions are achieved, and identify circumstances in which the benefits of more precise predictions can help justify new data collection and model development efforts.

For future efforts in Lake Erie and in other Great Lakes, a planned, deliberate, and coordinated effort should be made to characterize and present phosphorus load estimate uncertainty in a manner that is as consistent and clear as possible across all model types and for the consumption

of a wide audience of decision-makers and interested stakeholders. This includes the evaluation of statistical performance metrics and load-response figures with prediction intervals as discussed above. To better communicate uncertainty, it would be helpful to present different prediction intervals for mean model predictions, including but not limited to the 95% prediction interval shown in Figures 11-13 of the modeling report. There is no science-based or project-specific "requirement" to use a 95% prediction interval to characterize uncertainty in phosphorus load estimates. Prediction intervals associated with other levels of confidence (e.g. 90%, 80%) may also be appropriate based on project-specific needs would provide greater transparency in how the range of load estimates varies as a function of the level of confidence desired by decision-makers. This idea is consistent with a Data Quality Objectives type approach for evaluating the quality of information used for decision-making. One example is the use of a 70% prediction interval (and a 95% interval) in Figure 2a of Scavia et al. (in review, 70% prediction interval attributed to Stumpf et al.).

Dr. Emma Rosi-Marshall

Comments from Dr. Emma Rosi-Marshall, Cary Institute of Ecosystem Studies

1. Please comment on whether the evaluation of the models was adequate to inform how model results should be interpreted, given differences in model complexity and scale. Please identify any additional analyses that may be needed to improve future development and interpretation of the load-response curves for the eutrophication response indicators.

Comment: The use multiple models available to investigate the loading of P to Lake Erie is a robust approach. The models that are available allowed the team to develop investigate numerous indicators and the relationships to P loading. The fact that there are differences model structure, complexity, etc. among the 9 models makes this analysis more robust, as the outcomes of models are not reliant on the type of model used. The model output presented in figures 14-18 demonstrated clearly how the various models showed similar P loads that in relation to cyanobacterial blooms, hypoxic area and hypolimnetic DO concentrations. The models generally agree and suggest what P loading may be effective for reducing eutrophication. The nine models have all been used and peer reviewed in the past and they were selected because of their applicability and previous evaluations. I did not identify additional analyses to improve this approach; however, as new data and models become available, incorporating additional modeling efforts will continue to improve load estimates that are protective of water quality.

2. Please comment on whether the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives defined in the GLWQA (as reflected in Table 1 on page 7 of the document titled *Recommended Phosphorus Loading Targets for Lake Erie*).

Comment: The recommended targets outlined in the Recommended P Loading Targets for Lake Erie are based on multiple endpoints and are based on the modeling and long-term data on Lake Erie. These targets are appropriate based on this approach. The report concludes that adaptive management strategies should be implemented for the best future management of Lake Erie P loading. This is an excellent suggestion and I agree that implementing adaptive management would be advisable.

3. Please comment on whether scientifically-sound phosphorus load reduction recommendations to address *Cladophora* growth in the Eastern Basin of Lake Erie could be developed at this time.

Comment: Although there has been research on *Cladophora* (as indicated in the report this should be expanded to include other species of algae as well, e.g. expand to nuisance benthic algae, as suggested in the reports) in Lake Erie, it appears that there are some substantial knowledge gaps that remain and require some additional research. The report list some

additional new data to be collected such as long-term sampling of particular locations and/or use remote sensing to monitor nuisance benthic algae, but these approaches have their limitations. Sampling in specific locations is time consuming and by necessity is limited in spatial scope and could miss important blooms of benthic algae. It is unclear if remote sensing is at a point that it can be used with confidence, although these approaches are improving over time. That said, remote sensing coupled with high intensity sampling at a few sites seems like the best path forward for monitoring this spatially variable phenomena.

4. What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?

Comment: Including investigation of the other nutrients, e.g. N, was discussed in the report. Although P is a very important limiting nutrient, N dynamics should also be considered. That said, the long-term data on N in the Maumee River do not show changes in the long-term record. However, continued monitoring of N, suspended solids, zebra mussels, etc, should occur in the future to ensure effective long-term management of Lake Erie. Some aspects of N dynamics that should be continued monitored include inputs from various tributaries, especially considering the extensive agriculture in the watershed of Lake Erie. In addition, the influence of N dynamics on nuisance benthic algae (i.e. Cladophora and other species) should be investigated to increase understanding of the causes of these benthic algal blooms. The cycling of N in the system including investigating the various forms of N (e.g. nitrate, ammonium, dissolved organic N) and rates of N transformations should be considered for monitoring and research.

5. Please comment on the use of FWMC and any other approaches that should be considered to account for inter-annual variability in hydrology in assessing progress in reducing tributary loadings of phosphorus to the Lake.

No comment at this time

6. Please comment on the value of applying the existing eutrophication models on an ongoing basis to periodically evaluate phosphorus loading targets and eutrophication response indicators. What key elements should be included in the adaptive management approach to successfully implement and evaluate our nutrient reduction goals for Lake Erie?

This is highly recommended and would be essential for effective adaptive monitoring of the controls on P loading and responses in Lake Erie, especially considering the dramatic effects of invasive species, e.g. zebra mussels. Future invasions, climate change and other unforeseen changes in the ecosystem need to be monitored to allow for management to adapt over time. Key elements in adaptive management include long-term monitoring of nutrient inputs from tributaries and concentrations of nutrients in the lake. These activities should continue to be funded as these data are necessary for long-term management. Monitoring of eutrophication

indicators should continue and research funds should be allocated to investigate additional indicators of eutrophication. Research into new remote sensing approaches should continue to be investigated. As new satellite imagery continues to progress, these approaches should be implemented in the Lake Erie basin. Finally, the modeling approaches used in the suite of models should adapt and evolve over time. As new modeling approaches become available, these should be incorporated into the suite of models being used to investigate P loading. It would be advisable to continue to fund research into new ways to model the influence of P on Lake Erie.

Dr. Eric P. Smith

Comments from EPSmith

Overall Comments

The researchers have used state of the art models to estimate TP load reductions that should reduce Cladophera biomass, improve dissolved oxygen and chlorophyll a. The models that are used vary from newly developed to older models that have been consistently improved and updated. Several of the model have been used for analysis at different lakes and waterbodies with adequate results. It is clear from the documents that the researchers have invested considerable time into trying to use available data to address the issue and evaluate the strength of the models. I think they have done an excellent job of connecting the available data with the models and the results are valuable for supporting management decisions.

My main overall quibble with the results is that although the authors try to make a case that the modeling results in "good fits" it is not clear to me that this is the strongly supported. [Note: it would have been useful if the bar was set prior to the modeling exercise.] I think the difficulties with the approach lie not with the models that are used but the link between models and data. The simpler models are constrained by the need to evaluate a total load on an annual scale; the total number of years are the constraint in this case. For more complex, multi-compartment models, there is difficulty obtaining adequate data for all elements of the model. Hence, parameters that may be sensitive do not have the data support that is needed and the result is high uncertainty in the load reduction estimates.

One valuable result that I think could come from the overall exercise is a stronger statement about how to better link the monitoring programs to the models that are identified as "best" to address the relevant management questions.

Some questions and thoughts

Page 34: what is the right way to measures sensitivity in a statistical model – they use holdout of one observation.

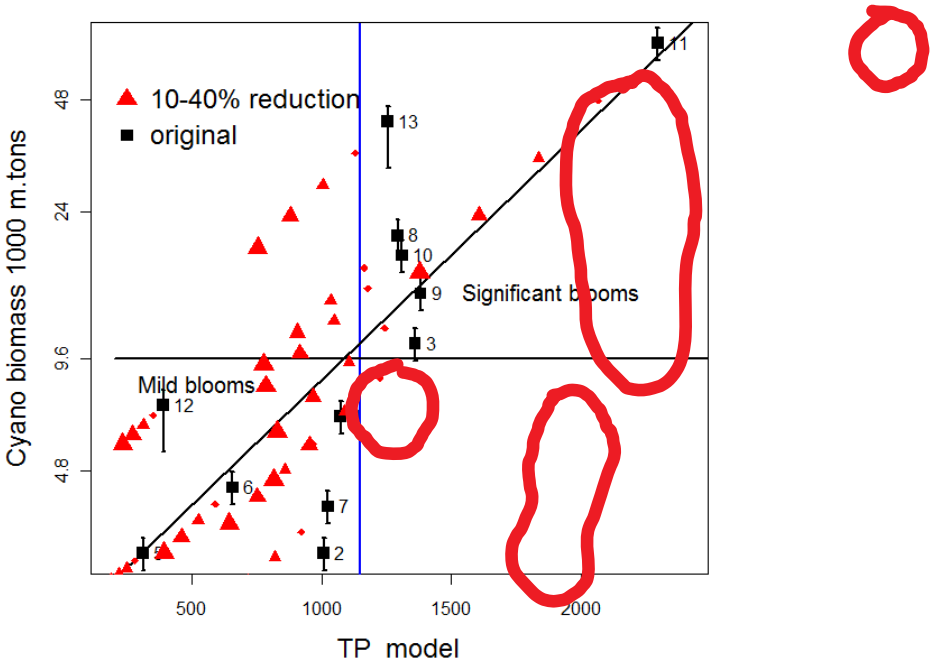
One does not have a clear sense of how large the sensitivities are. Is there a better way – give a measure. Not the discussion on the stat stuff mentions year 2011 and that estimates are within 90-105 of the unadjusted data. Why not compute an R^2?

B7-8 indicates that nutrients from Heidelberg for Maumee are released at a reasonable schedule but the releases from Michigan DEQ are lagged 5 years. I was confused by this and started to think about time scales and time frames associated with results. Does this affect information on the current effect of Detroit River? Would it be useful to have table to describe data used by different models, especially time scales and frames?

My opinion is that some results are slightly overstated. For example page 25: "Similarly, both statistical models (NOAA and UM/GLERL) were capable of resolving a substantial portion of the variability in annual bloom size by using Maumee River TP load as bloom predictor (R2 = 91.6 for the UM/GLERL model and R2=75.5 on the original, un-transformed scale for the NOAA model), confirming the relatively high predictive performance of both models."

Note that R^2 is not a measure of predictive performance rather it measures linear correlation. It is influenced by extreme observations (such as the data from year 2011), high leverage values (year 2012) and may not be relatively uncertain with small sample sizes. The "held-out" measures are more informative. The figure B1-2 illustrates the dilemma faced with the modeling – that Cynobacteria biomass varies considerably for a given model TP load. The data for 2011 was a blessing in that it will greatly increase R2 and reduce estimated uncertainty. However because it is extreme results need to be interpreted with some caution.

My view of the model results from page 25 are somewhat weaker and I have illustrated some concerns on the graph. Here, the model for the TP and CI relationship seems rather weak as a predictive model although the correlation is moderately high. This can be seen by the considerable variation in CI when TP is roughly constant – for example around 1200, the range in CI is roughly 10-45! This seems considerable and suggests something is missing from the scientific understanding of the processes that drive Cladophera production. Perhaps the results are adequate for management purposes but I do not think strong statements can be made about the relationships.



Apppendix B – I have also made a number of comments directly on the document

1. Please clarify the difference between SRP and DRP, if there is one. If they are the same, should only one be used in the text?

2. I think that more needs to be said (or modeled) about resuspension? I am of the opinion (perhaps incorrectly) that even if loads are set to zero, there is sufficient TP to create blooms for a considerable time. It might be worth noting that even with load reduction there may still be blooms.

3. Is spatial correlation included in the statistical models (or any other model) and calculation of uncertainty?

4. For the Bayesian models it might be useful to report on MCMC details: convergence, number of chains, thinning, etc.

The questions posed to the committee:

1. Please comment on whether the evaluation of the models was adequate to inform how model results should be interpreted, given differences in model complexity and scale. Please identify any additional analyses that may be needed to improve future development and interpretation of the load-response curves for the eutrophication response indicators.

    a. What data is needed to reduce the uncertainty?

    b. Can the TP models be linked to land use/runoff models to improve estimation of nonpoint contributions (identify spatial sources) and help target management actions for reductions.

    c. It might be useful to evaluate loading reduction based on the simple model proposed by Vollenweider (1976) as a simple approach. TPconc is the desired concentration of TP and CritTPload is the loading.

2. Please comment on whether the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives defined in the GLWQA (as reflected in Table 1 on page 7 of the document titled *Recommended Phosphorus Loading Targets for Lake Erie*).

    a. Although it seems the models suggest and immediate reduction in Cladophera, one must note that reductions in TP, if met, will likely not result in immediate Cladophera reductions. The graphs in Fastner et al 2016 (Combating cyanobacterial proliferation by avoiding or treating inflows with high P load—experiences from eight case studies, Aquatic Ecology) suggest that reductions might be required for up to 20 years before there is clear reduction in Cladophera.

3. Please comment on whether scientifically-sound phosphorus load reduction recommendations to address *Cladophora* growth in the Eastern Basin of Lake Erie could be developed at this time.

    a. Yes, although I think the uncertainty is moderate, especially since it is linked to nonpoint sources. The research group have a good understanding of the sources of loading and recommendations are sound. The adaptive management approach should support further decision making and adjustments to the load reduction recommendations.

    b. The paper by Mantzouki et al 2016 (Understanding the key ecological traits of cyanobacteriaas a basis for their management and control in changing lakes, Aquatic Ecology) might help improve models predicting effects of TP reduction.

4. What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?

    a. Perhaps one question that might be easily addressed is whether the reductions in TP will also result in reductions in TN. As noted in the Fastner et al. (2016) paper "Instead, the recovery of the lakes presented here and that of many other long-term case studies demonstrate that it is sufficient to reduce phosphorus concentrations below a specific threshold in order to reduce phytoplankton biomass, i.e. without the additional need to reduce nitrogen."

5. Please comment on the use of FWMC and any other approaches that should be considered to account for inter-annual variability in hydrology in assessing progress in reducing tributary loadings of phosphorus to the Lake.

    a. Models will have to account for changes in patterns in rainfall and snowmelt that are expected under climate change scenarios and evaluate if this will be a factor in Cladophera productivity.

6. Please comment on the value of applying the existing eutrophication models on an ongoing basis to periodically evaluate phosphorus loading targets and eutrophication response indicators. What key elements should be included in the adaptive management approach to successfully implement and evaluate our nutrient reduction goals for Lake Erie?

    a. I think it would be wise to invest some thinking into where and when to monitor and how to better connect monitoring and modeling. Better understanding of what initiates and drives the blooms is key to their reduction. Reducing uncertainty is key to improving estimated load reductions and improving model performance. Relating

model results to best management practices to reduce non-point source contributions is also key.

Dr. William Stubblefield

Phosphorus load Reduction targets for Lake Erie

Charge Questions: W Stubblefield Preliminary Responses

1. Please comment on whether the evaluation of the models was adequate to inform how model results should be interpreted, given differences in model complexity and scale. Please identify any additional analyses that may be needed to improve future development and interpretation of the load-response curves for the eutrophication response indicators.

   It would appear that the authors who developed the recommended phosphorus loading targets for Lake Erie did an excellent job of reviewing and considering the most appropriate and useful models available for evaluating nutrient inputs into Lake Erie. The pros and cons of each of the models were considered and the best aspects of each model used in developing the phosphorus loading recommendations. In many cases shortcomings were identified for the specific models in these highlighted as a source of model uncertainty.

2. Please comment on whether the recommended targets reflect the best available information on the drivers of cyanobacteria growth and seasonal hypoxia in Lake Erie and are appropriate to meet the nutrient Lake Ecosystem Objectives defined in the GLWQA (as reflected in Table 1 on page 7 of the document titled *Recommended Phosphorus Loading Targets for Lake Erie*).

   The recommended targets do seem to reflect the best available information and the state-of-the-science regarding the drivers of cyanobacteria growth and hypoxia in large lentic systems. Other nutrients such as nitrogen levels or micronutrient concentrations may be appropriate but little is known about those in the scope of the types of models being employed.

3. Please comment on whether scientifically-sound phosphorus load reduction recommendations to address *Cladophora* growth in the Eastern Basin of Lake Erie could be developed at this time.

   The authors acknowledge that there are difficulties relating *Cladophora* biomass to Lake Erie phosphorus loads and that appropriate, validated population models are not available at this juncture. A variety of concerns were identified but the Objectives and Targets Test Team was not prepared to establish a loading target for *Cladophora* in the Eastern basin of Lake Erie. A number of recommendations for continued monitoring programs and additional research were provided and it was suggested that the phosphorus loading recommendations for the remainder of Lake Erie would likely help control *Cladophora* in the Eastern basin of Lake Erie.

4. What recommendations can the SAB provide for development of an approach to help determine whether consideration of nitrogen control, in addition to phosphorus, is warranted in Lake Erie to prevent harmful algae blooms and manage hypoxia? In particular, what questions, relationships, or research priorities related to nitrogen loading (different forms and sources) and in-lake cycling must be addressed?

   Previous research in Lake Erie, and other lentic systems, has suggested that algal populations may not be controlled entirely by phosphorus concentrations, but may also be associated with nitrogen concentrations as well as micronutrient concentrations. Consideration of these parameters may reduce uncertainty and variability in the nutrient models that are been applied to the Lake Erie issue.

5. Please comment on the use of FWMC and any other approaches that should be considered to account for inter-annual variability in hydrology in assessing progress in reducing tributary loadings of phosphorus to the Lake.

   The use of flow-weighted mean concentrations (FWMC) are appropriate as a means to account for inter-annual variability in nutrient concentrations. However, it is important to consider the averaging period over which the FWMC is calculated. It is important that the best estimate of nutrient exposure concentrations is provided and this related to algal population response. Alternative approaches may be appropriate and should be considered as a means to reduce model uncertainty.

6. Please comment on the value of applying the existing eutrophication models on an ongoing basis to periodically evaluate phosphorus loading targets and eutrophication response indicators. What key elements should be included in the adaptive management approach to successfully implement and evaluate our nutrient reduction goals for Lake Erie?

   Application of the existing eutrophication models on an ongoing basis is appropriate and will help address questions associated with model uncertainty and model validity. Implementation of an ongoing monitoring program will provide the data, in the future, to address modeling concerns and will allow us to address questions associated with those areas where we are currently unsure about model validity, e.g. nearshore nutrient concentrations.