Science Advisory Board
Nutrient Criteria Review Panel

Individual Pre-Meeting Comments

Preliminary Evaluation of *Methods and Approaches for Deriving Numeric Criteria for Nitrogen/Phosphorus Pollution in Florida's Estuaries, Coastal Waters, and Southern Inland Flowing Waters*.

Developed in Preparation for the December 13-14, 2010 Meeting of the Nutrient Criteria Review Panel.

Table of Contents

Dr. Boynton………………………...…………………………….. 1
Dr. Chapman ……………………………………………….... 5
Dr. David………………………………………………………12
Dr. Diaz ……………………………………………………..14
Dr. Giblin ....…......21
Dr. Meyer ……………………………………………………27
Dr. Noll …………………………………………………......32
Dr. Paerl ……………………………………………………. 33
Dr. Reckhow ……………………………………………...37
Dr. Schneider ……………………………………………..39
Dr. Sharpley …………………………………………………41
Dr. Steinman ……………………………………………...43
Dr. Solow ……………………………………………………51

Dr. Walter Boynton

2. Estuaries
    a. Are the data sources identified appropriate for use in deriving numeric criteria in Florida's estuaries? Is the SAB aware of additional available, reliable data that EPA should consider in delineating estuaries or deriving criteria for estuarine waters? Please identify the additional data sources.

At this initial stage of panel work this is a question I can not answer in any depth. I will say that the report lays out quite a number of data sets…the list is impressive. I think discussions will raise questions about other possible data sets. My impression is that the EPA folks have done a good job of identifying and organizing data sets. There may be some data sets that, in one way or another, have been "disqualified" and these should be given a second look. In specific estuaries these data sets may be of great utility.

    b. Are the assessment endpoints identified in Sections 2.3 and 3.2 (healthy seagrass communities; balanced phytoplankton biomass and production; balanced faunal communities) appropriate to translate Florida's narrative nutrient criterion into numeric criteria for Florida's estuaries, given currently available data? Does the SAB suggest modifications or additions to these endpoints? A literature review of endpoints considered can be found in Appendix B.

Again, at this early stage it is a bit hard to tell. They do seem reasonable to me because these sorts of data are widely available and the current state of knowledge supports these as reasonable indicators of system function. One of the endpoints includes phytoplankton production. I did not see anything about these rate measurements in the text. Are these to be modeled from biomass, nutrient and light data? Are there enough measurements of production in Florida to support a modeling effort? Also, hypoxia is an issue mentioned in several areas. Are there data concerning water column and sediment rates of oxygen utilization? If not, how will models be calibrated and verified? I thought the literature review was solid and represents a great deal of work.

    c. EPA describes potential approaches in Section 3.3 (reference conditions, stressor response relationships, and water quality simulation models) for deriving numeric criteria in Florida's estuaries. Compare and contrast the ability of each approach to ensure the attainment and maintenance of natural populations of aquatic flora and fauna for different types of estuaries, given currently available data?

This is a very large and complicated question. I think discussions (and I believe testimony from Florida folks) will shed some useful light on this question. At this point I think the reference condition approach is the most straightforward if areas can be identified having the required characteristics. That might be difficult to do. Historical data (representing pre-nutrient impact conditions)seem to be sparce. The modeling approach is by far the most sophisticated, encompassing and difficult. It has advantages of directly linking nutrient sources (to be managed) with effects through some complicated processes; experiments can be readily conducted with the models (and all the other things models can do). But, there are serious risks associated with complex modeling including the need for a lot of data, time, costs and there is no guarantee the thing will work well enough to be useful…model world is not necessarily closely related to the real world to be managed. The statistical approaches offer a more rapid (likely), realistic (based on observation) approach. But, there is the need for strong relationships among key *variables to emerge…this may or may not happen. I trust that discussions with other panel members and presentations will help focus recommendations in this area.*

Comments from Report with Emphasis on Estuaries:

I have completed an initial reading of the Report titled "Methods and Approaches for Deriving Numeric Criteria for Nitrogen/Phosphorus Pollution in Florida's Estuaries, Coastal Waters, and Southern Inland Flowing Waters" (November 17, 2010). I will need to re-read and think more about several sections of this report before reaching more than tentative conclusions. I look forward to further discussion at the EPA Panel meeting (Dec 13-14, 2010). I have listed below some comments based on my initial reading of this document (and a more superficial reading of the Appendix documents).

1. Pg 18. It would be useful to have some side-bar boxes with definitions of the many important terms that appear in the Executive Summary. I know these items are defined elsewhere…for the sake of rapid and accurate understanding, some side-bar definitions would be helpful
2. I think the WQ criteria specific to sub-regions is a good concept because it directly addresses the idea that these estuarine ecosystems are not identical…special circumstances require some criteria flexibility.
3. Suggested endpoints seem quite reasonable….some are certainly more readily evaluated than others. More on this later.
4. The three approaches being considered also seem reasonable. They are however quite distinct in terms of the amount of effort required to reach a conclusion, the amount of data needed, the certainty associated with results, the detail with which results can be used to correct pollution sources, and likely many other issues.
5. I was pleased to see attention and considerable detail provided relative to downstream water quality standards (40 CFR 131.10(b)). This is a challenging issue but quite central, in my opinion, to a successful pollution control program. In at least one area of the country, river segments are being de-listed because local criteria are being met, but these same segments are transporting large amounts of nutrients that most certainly have negative impacts on down-stream estuarine areas.
6. pg. 24. The fact that there are a very large number of estuaries (120) will make this a challenging issue. While sharing many characteristics there will be particular features associated with these systems. Most are quite small (average area ~5 mi^2).
7. Pg 25. Effects of N and P pollution should also include food web changes. There is some redundancy in the primary producer listing and that might be improved. N and P effects also change the mix of primary producers and the list of producers should include phytoplankton, SAV (and epiphytes), macro-algae and benthic microalgae.
8. Pg. 26. The diagram should include in the LOW OXYGEN box a mention of changed nutrient biogeochemical pathways and the positive feedbacks these changes have on the progression of eutrophication.
9. It seems that the use of nutrient concentrations and nutrient loads are used clearly at times and at other times it seems that a decision has not yet been made as to how these related variables will be used (Pg. 27). Am I missing something here?
10. I'm glad to see that both N and P are being considered in this evaluation. Figure 2.1 nicely summarizes a great deal of material.
11. Useful summary of assessment endpoints and indicator variables in Tables 2-1 and 2-2.

12. Pg. 38. The relationship between chlorophyll-a and TN and TP in the text seems assured. That might not be the case in all these systems. The relationship is likely a bit more complicated. I'm sure the authors are very aware of this. Some scaling of causative variables might prove useful, as limnologists have found. Same comment concerning DO and TN and TP.
13. Pg. 39. The indicator variables (TN, TP and chlorophyll-a) seem well justified. The discussions on pages 39-40 are clear and useful in this context. The endpoints not selected also seem justified. It would be prudent to keep an eye on some or all of these as nutrient reductions (hopefully reductions) begin to have effects. In other words, some of the non-selected endpoints might be very important in some estuarine systems and less so in others (pg 40).
14. Descriptions of data sources seem clear and through (pg 41-43).
15. Pg 45. I think there are real advantages to recognition that sub-system numeric criteria may be needed. The Chesapeake Bay has several DO criteria some of which have a living resource basis and some have a basis in historical measurements.
16. Good points relative to relationships between TN and TP and indicator variables.
17. Pg 49. Water clarity is influenced by chlorophyll which is generated by N and P loadings…the N and P by themselves do not create turbid conditions. Sediments and CDOM should also be mentioned.
18. Pg. 49. I do not see how chlorophyll-a will tell much about species of phytoplankton. HPLC analysis could be of some utility although I'd be surprised if such data are generally available.
19. Pg 50. The approach summary table is useful. I think the regression approach is sold a bit short (more complex models are available). To be fair, I think the simulation approach should also include an indication they take a lot of time to develop and cost a good bit of cash….that's my experience. In this same category, how will a simulation model be calibrated and verified if there is little site data available…that seems like a weakness rather than a strength.
20. It has been quite awhile since I was reasonable familiar with Florida waters. So, the next comment may be off base. I think it will be difficult to use the reference approach because there are few historical observations (i.e., when systems less impacted). TN and TP data sets seem recent. There are some chlorophyll observations from the past, but not many.
21. I do hope the work of Steward and Lowe is given some very serious attention. Results are quite striking to say the least.
22. There is considerable discussion concerning use of simulation models. I urge some caution here. The description sounds great and there are some issues that we can only address with simulation. However, this would be a huge job, take considerable time (note the Chesapeake Bay model has been under development for about 25 years and still does not predict inter-annual hypoxia volumes well), and useful results are not guaranteed. So, be careful! Has the decision already been made to employ coupled models as the primary tool?
23. Pg. 65. In the water clarity/SAV discussion the effects of epiphytic growth on SAV leaves should be considered…in some areas this is as important as water column light attenuation.
24. Pg 66 (bottom). Is there any though being given to generating new data in areas that are data poor? Seems necessary with any of the approaches being considered.
25. Pg 67-69. Most of the SAV coverage data are from the last 2 decades. Some data from the 1950s. Does the more recent data mainly tell us about distributions after there have been declines? How is a "healthy state" determined if there is a very limited view of the past?
26. Pg 71. If WQ targets can not be met without reducing TN and TP loading below natural levels….what does this mean? Need some additional clarity here.
27. Pg 74. Is it worth considering point source loads as small as 0.1 mgd?
28. Pg 74. Is all the stream reach detail supported by measurements? Is enough known to justify this level of detail?

29. Pg 76. This is a general comment on the preceding pages. The evaluation period seems quite short…does it consider wet and dry periods, for example? Second, is there enough data to justify and support the level of detail indicated here?
30. Pg 79. I'm very surprised that there are so few chlorophyll observations available for coastal waters. Is this really true?
31. Pg. 83-84. The degree to which chlorophyll is predicted is quite low (R2 = 0.52). Is this satisfactory? There were only 62 observations available? Does this work need to be extended?
32. Pg. 97. What is the level of detection for TP?
33. Pg. 109-113. I think the stress-response discussion is too brief and more could be done…perhaps a great deal more. For example, why not investigate more complex statistical models…as limnologists have done. What about "scaling" the x-variable in these analyses. It is heartening that some variability is explained with the simplest of models. I'm not a statistician so my observations here are of a qualitative nature. But, I think more could be done here.
34. I think the issue of DPVs is especially important. I'd like to hear more about this aspect of the effort. I also liked the issue of local spatial conditions being recognized. I trust this will avoid crazy results.
35. Pg. 126. I'm not sure why the south Florida presents such a problem. With all the management done there (and the associated data bases) it would almost seem this would be the place to do the best work. I'd like to hear more about issues here.

Dr. Piers Chapman

As populations expand, so the pressure on the natural environment continues to increase. One of the most important areas of concern is that of water pollution, in particular pollution from the nutrients nitrogen and phosphorus. These have been shown to cause major changes in many areas (e.g., Diaz, 2001; Diaz and Rosenberg, 2008), with effects such as algal blooms, hypoxia/anoxia, fish kills, and overall changes in trophic levels. The State of Florida already has a narrative criterion for water quality standards, which states (in part) that:

> in no case shall nutrient concentrations of a body of water be altered so as to cause an imbalance in natural populations of aquatic flora or fauna (F.A.C. 62-302-530 (47)(b)).

Additional criteria exist for potable water supplies in Florida, where the drinking water maximum for nitrate has been set at 10 mg/l (it is unclear from the draft document whether this applies to nitrate-nitrogen or to the concentration of nitrate ion; there is a fourfold difference between the two). In 2009, the state published a "State of Florida Numeric Nutrient Concentration Development Plan" (FEDP, 2009), which contained EPA-generated and recommended criteria for inland waters within the state (Table 1). There are currently no criteria for estuaries or coastal waters, which extend out to three nautical miles from the coastline.

Table 1. EPA-generated nutrient criteria recommended for Florida's ecoregions. Region IX comprises the northern part of the panhandle, away from the coastal strip. Region XIII comprises the southern portion of the state essentially occupied by the Everglades, and Region XII comprises the remainder, including the panhandle coastal strip.

Lakes and reservoirs

Parameter	Region IX	Region XII	Region XIII
TP (µg/l)	20	10	17.5
TN (mg/l)	0.36	0.52	1.27
Chlorophyll a (µg/l)	4.93	2.6	12.35
Secchi depth (m)	1.53	2.1	0.79

Rivers and streams

Parameter	Region IX	Region XII	Region XIII
TP (µg/l)	36.56	40	NA
TN (mg/l)	0.69	0.9	NA
Chlorophyll a (µg/l)	0.93	0.40	NA
Turbidity (NTU)	5.7	1.9	NA

The draft document we have been asked to review considers methodologies by which EPA can determine suitable numeric criteria for Florida's estuarine and coastal waters, and the inland flowing waters within the southern part of the state. In my opinion, the authors have done a good job of defining potential statistical methods that could be used to set numeric criteria. I do have some objections, however, to some of the assumptions they make about the relationships that will form the basis for these statistics. Given the short period we were allowed to read the draft document and provide this review, my response is shorter than it could be.

Clearly, as stated by Paerl (2009), there are many differences between fresh and estuarine and coastal waters. The most obvious is that there is no requirement for estuarine and coastal waters to met drinking water standards because the increased salinity renders them unfit for drinking. There is also a generally applicable difference between freshwaters and saline waters in that phytoplankton production in the former tends to be controlled mainly by phosphorus, while in more saline regions nitrogen is generally considered to be limiting (see e.g., Paerl, 2009 and references therein). Obviously there are exceptions to this, such as the infrequent phosphate limitation on phytoplankton productivity off the Mississippi (Sylvan et al., 2006), but by and large it is nitrogen addition rather than phosphorus that induces phytoplankton growth in coastal regimes, and has resulted in eutrophication and the annual development of deleterious hypoxic and/or anoxic zones in Chesapeake Bay (e.g., Boynton and Kemp, 2000) and in the northern Gulf of Mexico (e.g., Rabalais et al., 2007). However, it is clear that if numeric criteria are set, they need to include both nitrogen and phosphorus (see arguments in Paerl, 2009).

We have been given six charges to consider. I take these in order.

Charge questions

1. <u>General approach</u>

The general approach is laid out in Chapter 2 of the draft document. I believe that the overall procedure as laid out in paragraph 2.1.2 is adequate to produce the desired result, provided the final bullet in item 7 is adhered to. This demands that criteria set for upstream regions must not result in overwhelming downstream regions (see Charge 6). However, I have a problem with the idea that there is a direct relationship between criteria set for TN and TP and the expected criteria assumed for chlorophyll a, as suggested in Fig. 2-1, since in many instances, depending on season, temperature, or some other parameter, there is no clear relationship. I will discuss this, as well as potential issues with some of the conclusions stated in Table 2-2, further under Charge 3.
This affects both the selection of potential assessment endpoints and indicator variables discussed in Section 2.3; for example, the document does not define what is meant by “total nitrogen (TN)” or “total phosphorus (TP). This has major impacts on what needs to be measured, especially for nitrogen species.

Regarding the question of whether the categories of water bodies selected are suitable, I have some reservations about separating estuaries from coastal regions. The reason for this is that there may be considerable water exchange between the two. While for much of the year such exchange may be limited, when frontal systems move across the state or hurricanes strike, they are generally accompanied by large amounts of rain that will increase the exchange rate between freshwater, estuarine and coastal areas. I saw no discussion in the draft document about dealing with such infrequent, but likely high-impact, occurrences. Presumably this will be dealt with in terms of setting percentile values that may be exceeded on a small number of occasions. However, it does seem reasonable to separate the coastal waters of south Florida from the eastern and western coasts, because of the very different nutrient requirement of the coral reefs. Using a single standard for all Florida coastal waters would likely lead to criteria that are too high in the southern region.

2. <u>Florida estuaries</u>

The draft document states that EPA is considering developing estuarine nutrient criteria on a system-specific basis (p.45). I general, this is a good idea for the reasons given in the document, especially the variable rates of development in the estuaries’ watersheds. The U.K. in the 1970s adopted water management districts that encompassed the complete watershed of major river systems, including estuaries, and this proved to be a very successful way to manage water pollution issues. Of course, estuarine conditions in the U.K. and Florida are vastly different because of parameters such as water flow rates, temperature, tidal range, offshore current velocities etc., but that does not invalidate this approach.

If seagrass community structure and health can be monitored based on depth of colonization and water clarity (p. 49), then this would seem a relatively simple and cheap way to monitor estuarine health. While not an expert on seagrasses, I understand that populations can and do vary by large amounts from year to year for reasons that are not presently understood. However, the use of a water clarity metric should not be disqualified for this reason, and it would seem reasonable to consider water clarity as a good measurement of community structure health even for estuaries where seagrasses are not found naturally. I do not see, however, how chlorophyll a concentrations by themselves will have any direct relationship with changes in phytoplankton species composition; one can have large blooms of mixed phytoplankton in response to changes in nutrient concentrations, as is found, for example, in upwelling systems (e.g., Olivieri 1983)

Regarding the question of potential approaches to derive numeric nutrient criteria (section 3.3), reference conditions will certainly work where one has a long-enough data set to be able to show conclusively that a water body has become impaired over time. However, as stated on p. 51, it appears that there are no "impaired estuaries [in Florida] for which historical water quality data could provide a suitable reference condition." It is not clear to me if this means that there are no impaired estuaries in Florida, or merely that the datasets are too short to be able to define nutrient concentrations for when the estuaries were unimpaired. Given the times listed in Table 3-3, I suspect the latter is true, in which case how do you define an "unimpaired state"?

I am not familiar with any of the models discussed in section 3.3.3, although the results from all of them will depend on the amount and quality of the data used to set them up. In this instance, similar problems to those given above regarding establishing suitable reference conditions apply. Again, these models assume a close relationship between nutrient inputs and chlorophyll a concentration, which may not be justified (see comments on Charge 3). However, if it can be shown that several Florida estuaries share common features, and that the models can be validated, then there should be a large reduction in the amount of work required to set numeric nutrient standards. The example given for Pensacola Bay is clear, but because there is no attempt to measure non-point sources of nutrients, I am not sure these are dealt with.

3. <u>Florida coastal waters</u>

While I agree that it is relatively easy to monitor chlorophyll a concentrations by means of satellites, I do have a problem with the idea that chlorophyll a concentration relates directly to nutrient concentration and can therefore be used as a proxy to set nutrient criteria for estuarine and coastal water bodies. The paper by Walker and Rabalais (2006), which is cited in the draft document, only found about 40% of the variance in phytoplankton production could be ascribed to nutrient concentration, and this was in an area of the northern Gulf of Mexico known to be affected strongly by nutrient inputs from the Mississippi River (in this case, they had to lag the chlorophyll data by one month because of the time taken for nutrient-enriched river water to travel to the observation site in the Gulf from its monitoring site above Baton Rouge). The main reasons for frequently low r^2 value in coastal waters for correlations between chlorophyll a and TN or TP are:

i. That there is no consistent relation between chlorophyll a content and nutrient uptake by phytoplankton. As stated in section 5.4.1.2, "although the underlying relationship certainly exists, in nature it has often been difficult to quantify because of the complexity of nutrient dynamics, the influence of interfering factors, and the temporal and/or spatial scales of the data and analyses."

As an example, there is often a major seasonal effect on phytoplankton growth rates that results from changes in water temperature. Because of this, winter nutrient concentrations off the Mississippi delta can be very high but not accompanied by increased chlorophyll a concentrations. Similarly, the well-known Redfield ratio (Redfield et al., 1963), used to link nutrient uptake to oxygen depletion and carbon production, refers specifically to carbon, not chlorophyll a, and it is known that the carbon: chlorophyll a ratio in phytoplankton varies by a factor of at least five, and possibly by an order of magnitude, depending on the physiological state of the plankton (Banse, 1977).

ii. That maximum chlorophyll concentrations do not necessarily develop at the site of nutrient supply, but downstream from it because of tidal and other current action. Thus, high chlorophyll a concentrations do not mean high nutrient concentrations at the site of the measurements. As an example, effluent from the Mississippi River plume can be observed many miles from its source by examining satellite imagery of chlorophyll a concentrations (Walker et al., 1994).

This argument also applies to the potential use of dissolved oxygen concentrations to back-calculate nutrient concentrations. Florida has oxygen standards for estuarine and coastal waters such that the average DO shall not be less than 5.0 mg/l (~3.5 ml/l) in any 24-hour period, and shall never be less than 4.0 mg/l (~2.8 ml/l) with normal daily and seasonal fluctuations maintained (draft document, p. 38). Because the decomposition of the organic matter associated with high chlorophyll concentrations will take time (on the order of days to weeks), low oxygen measured at a particular site need not have been produced there, but could have moved under the influence of bottom currents. However, because the western shelf of Florida is very wide, this may not be a problem except near the mouths of estuaries subject to nutrient input.

iii. That there is exchange between the water further off the coast than 3 nautical miles with that in the defined coastal zone. While much of the coastal regions along both the east and west coasts of Florida are generally affected by low nutrient surface water from the Loop Current and Gulf Stream, high-nutrient water can be either advected into the region or upwelled. On the western side of the state, cold water may be upwelled at De Soto Canyon, particularly if a cold-core edge eddy from the Loop Current is trapped there (Belabbassi et al., 2005). This water may then be moved onshore along the Florida panhandle depending on the local wind field. A similar effect may also occur following frontal systems or hurricanes; in both cases high-nutrient runoff from rivers west of Florida is to be expected, and in summer the prevailing winds are generally westerly (see Schmitz et al., 2005, and references therein), which again would bring the high nutrient water into Florida coastal areas. On the Atlantic coast, kinematic upwelling along the edge of the shelf can bring nutrient-enriched water to the surface, and this again can spill onto the shelf (Atkinson, 1985)

The same argument (p. 82) is used for coastal waters as for estuaries, viz. that "EPA has not yet identified any impaired estuaries [in Florida?] for which historical water quality data could provide a suitable reference condition." As stated above, it is unclear quite what this statement means.

At present, it certainly seems from the data presented in Fig. 4-9 that chlorophyll a concentrations along the east, west, and northwest coasts of Florida are generally low, with 75 percentile values of about 2, 3.5 and 1.5 µg/l respectively, so I may be overemphasizing the potential problems. These low concentrations suggest that nutrient concentrations are similarly low, even allowing for large variations in carbon: chlorophyll ratios. However, as stated on p. 90, the status of future satellite missions that could monitor chlorophyll a concentrations remains uncertain.

4. <u>South Florida inland flowing waters</u>

It appears that EPA has set a standard for total phosphorus, but not nitrogen, in the Everglades Protection Area. This standard, 10 ppb, equates to approximately 0.3 µM of phosphorus, presumably measured as phosphate. This is a very low figure for natural fresh waters, and suggests that the ecosystem is used to surviving in conditions of phosphate deprivation/ limitation. It seems likely that either the reference condition approach or the stressor-response approach should work for these waters, provided the necessary data can be collected. In the reference condition approach, repeated surveys of invertebrates will show changes in community structure and diversity, but this is a time-consuming methodology, particularly if many sites need to be sampled regularly. The stressor-response approach should also work if a suitable relationship between chlorophyll a and nutrient load can be demonstrated (p. 103), but several of the same caveats apply here as for setting limits in coastal waters. Selecting "least disturbed sites" also does not appear to be easy in this region that has been subject to active management for many years.

5. <u>South Florida marine waters</u>

Given the large database for South Florida marine waters obtained through the SERC water quality monitoring network, it would appear that this is one area where definitive relationships between nutrients and chlorophyll a should work, or at least be observable. Again, nutrient and chlorophyll a concentrations in these waters are very low, and need to remain in this state if the coral reefs along the Florida Keys are to remain healthy (although nutrients per se do not cause coral dieoff, rather it is the reduced water clarity from increased algal production that is the problem (Gibson et al., 2009, Szmant, 2002). So the question remains as to why there is a need to use geometrically-adjusted data to define any relationships, given that the 10-90 percentile variability for TN is only about one order of magnitude, and considerably less for TP (Fig. 5-17).

The CDFs for TP and TN shown in Fig. 5-18 suggest that the 75% limit for TP is about 0.01 mg/l (approximately 0.3 µM), while that for TN is about 0.32 mg/l, equivalent to about 23 µM if measured as nitrogen. This TN value seems very high, being typical of values from upwelling regions. If, however, TN is measured as nitrate, then it equates to about 5 µM, which seems far more realistic for an oligotrophic region such as Biscayne Bay. The TN:TP ratio then changes from 70 to about 17, which is close to the generally-accepted Redfield ratio. How is TN defined in this case?

6. <u>Downstream protection values for Florida estuaries and South Florida marine waters</u>

It is exciting and refreshing to see a regulatory agency attempting to set nutrient criteria in the way described in Chapter 6. Most attempts start by setting criteria for upper reaches of streams and rivers, which, as stated here, mean that even greatly decreased limits can still overwhelm downstream water bodies. However, on p.123 it is stated that

> "For the purpose of computing DPVs, the estimate of the average protective loading would exclude loads resulting from direct atmospheric deposition to estuarine surface waters and point-source loads discharged directly to estuarine waters. Similarly, the estimate of total streamflow would not include freshwater inputs resulting from net deposition onto the surface of the estuary (i.e., precipitation minus evaporation) and point source discharges of freshwater into the estuary."

Presumably these inputs (where available, or some estimate of them) will need to be subtracted from the total allowable load in the estuary before applying equation 6-1, or there will likely be an underestimate of the total loading to a particular water body.

I agree with the statement on p.127 that the use of equation 6-6, where $L_1 = c_1q_1$ (from equation 6-4) is likely to give much better results than equation 6-5 for calculating DPVs for South Florida marine waters. The application given for Pensacola Bay seems consistent with what is suggested. The question, of course, is whether the bay is considered to be in a natural or impaired state.

Finally, the question arises as to what the agency will do in the event that nutrient concentrations in any of the water bodies discussed here exceed the established criteria, as will certainly happen. If the intention is to use some sort of percentile time (e.g., the 75% percentile), this will require large quantities of data, which may not yet exist, in order to establish robust numeric criteria.

References

Atkinson, L.P., 1985. Hydrography and nutrients of the southeastern U.S. continental shelf. In: *Oceanography of the Southeastern U.S. Continental Shelf*, ed. L.P. Atkinson, D.W. Menzel and K.A. Bush, American Geophysical Union, Washington, D.C., pp. 77-92.

Banse, K., 1977. Determining the carbon-to-chlorophyll ratio of natural phytoplankton. Marine Biology, 41, 199-212.

Belabbassi, L., Chapman, P., Nowlin, W.D., Jr., Jochens, A.E. and Biggs, D.C., 2005. Summertime nutrient supply to near-surface waters of the northeastern Gulf of Mexico: 1998, 1999, and 2000. Gulf of Mexico Science 23(2) 137-160.

Boynton, W.R. and Kemp, W.M., 2000. Influence of river flow and nutrient loads on selected ecosystems: a synthesis of Chesapeake Bay data. In: *Estuarine Science: A Synthetic Approach to Research and Practice*, ed. J.E. Hobbie, Island Press, Washington, D.C., pp 269-298.

Diaz, R.J., 2001. Overview of hypoxia around the world. J. Environmental Quality, 30, 275-281.

Diaz, R.J. and Rosenberg, R., 2008. Spreading dead zones and consequences for marine ecosystems. Science, 321, 929-929.

FDEP, 2009. State of Florida Numeric Nutrient Concentration Development Plan. Florida Department of Environmental Protection, Tallahassee, FL, 138 pp.

Gibson, P.J., Boyer, J.N. and Smith, N.P., 2009. Nutrient mass flux between Florida Bay and the Florida Keys National Marine Sanctuary. Estuaries and Coasts, 31, 21-32, doi: 10.1007/s12237-007-9032-5.

Olivieri, E.T., 1983. Colonization, adaptations and temporal changes in diversity and biomass of a phytoplankyton community in upwelled water off the Cape Peninsula, South Africa, in December 1979. S. African J. Marine Science, 1, 77-109.

Paerl, H.W., 2009. Controlling eutrophication along the freshwater-marine continuum: dual nutrient (N and P) reductions are essential. Estuaries and Coasts, 32, 593-601; doi: 10.1007/s12237-009-9158-8.

Rabalais N.N., Turner, R.E., Sen Gupta, B.K., Boesch, D.F., Chapman, P. and Murrell, M.C., 2007. Hypoxia in the northern Gulf of Mexico: does the science support the plan to reduce, mitigate, and control hypoxia? Estuaries and Coasts, 30, 753-772.

Redfield, A.C., Ketchum, B.H. and Richards, F.A., 1963. The influence of organisms on the composition of seawater. In: *The Sea*, ed M. Hill, Wiley Interscience, 2, 26-77.

Schmitz, W.J., Biggs, D.C., Lugo-Fernandez, A., Oey, L.-Y., and Sturges, W., 2005. A synopsis of the circulation in the Gulf of Mexico and on its continental margins. In: *Circulation in the Gulf of Mexico*, ed. W. Sturges and A. Lugo-Fernandez, American Geophysical Union, Washington, D.C., pp. 11-29.

Sylvan, J.B., Dortch, Q., Nelson, D.M., Maier Brown, A.F., Morrison, W. and Ammerman, J.W., 2006. Phosphorus limits phytoplankton growth on the Louisiana shelf during the period of hypoxia formation. Environmental Science and Technology 40, 7548-7553; doi: 10.1021/es061417t.

Szmant, A.M., 2002. Nutrient enrichment on coral reefs: is it a major cause of coral reef decline? Estuaries, 25, 743-766.

Walker, N.D., Fargion, G.S., Rouse, L.J., and Biggs, D.C., 1994. The Great Flood of summer 1993: Mississippi River discharge studied. EOS, Trans. AGU, 75, 409-415.

Walker, N.D. and Rabalais, N.N., 2006. Relationships among satellite chlorophyll a, river inputs, and hypoxia on the Louisiana continental shelf, Gulf of Mexico. Estuaries and Coasts, 29, 1081-1093.

Dr. Mark David

Much of this draft document is outside my area of expertise, as I have no experience with water quality in estuaries or coastal waters. However, I am familiar with development of nutrient criteria for flowing freshwaters, and some of the same principals apply. Nutrients (N and P) can cause problems through algal production and resulting effects on dissolved oxygen (low DO in early morning from diel cycling), as well as turbidity impacts, and from the decomposition of dead algal and plant biomass that leads to low DO. However, nutrients are just one impact on water quality, and often need to be put in the context of other impacts on the biotic integrity such as habitat degradation, sediment, light, and toxic elements. In many systems reducing nutrients might not be in the top few changes that would greatly improve the biological integrity of the water body. I didn't find much in the document that describes how these other controls on aquatic integrity might have impacts in addition to or beyond the effect of nutrients.

Clearly the best way to understand N and P impacts are site specific studies where all factors can be evaluated, and hopefully the role of nutrients can be measured or determined. On page 23 of the draft document EPA indicates that Florida's method of doing this was too "resource-intensive and time-consuming." EPA then states that this case-by-case interpretation was "insufficient to ensure protection of applicable designated uses." This then leads to EPA determining that more broadly defined numeric nutrient criteria would essentially get the job done more quickly, and provide the necessary protection. I do agree in the perfect world with unlimited dollars to reduce nutrients this would be an approach that would move more rapidly. However, to reduce nutrients in many systems will involve reducing losses from agricultural areas, which are often more important than point sources. In many cases, we don't have very good methods to accomplish this to the degree needed to make major reductions in N and P concentrations in streams and further downstream. Even when TMDL's are completed in agricultural areas, they have no led to changes in agricultural losses, only in NPDES permits when they are present. The overall result is that the TMDL is not met. That is where a more site specific approach might be useful. I would assume Florida would first look at streams and coastal waters with known problems, understand the relationship with N and P, and determine how to reduce N and P concentrations when they are shown to be the major problem. Therefore, dollars could be targeted towards the worse problems, rather than attempting to reduce everything at once, something that is unlikely to occur.

To develop the nutrient criteria EPA proposes to use a variety of methods, because for most systems the detailed knowledge of what is effecting the biology may not be known, and the exact role of N and P is not known as well. Therefore reference conditions, modeling, and stressor-response relationships are proposed. Each can work, but all have many problems in that the relationships are known to be fairly complex (hence the need for site-specific understanding), so that strong and convincing regression or modeling results are often not obtained. There is also the view in the stressor-response relationships that if a relationship is found between chl-a and N or P (no matter that it may be weak), with no evidence of other impacts or relationships, that is enough to set the standard. I question whether that is an adequate stressor-response relationship to set a standard.

In some of the modeling described to set criteria in chapter 3, it seems that nonpoint source loadings are not well incorporated (e.g., page 75). The LSPC model describes “wash-off of water quality constitutes using simple relationships” with accumulation and wash-off rates playing an important role. I am not familiar with this model and it is not one commonly used to model nutrient losses from agricultural sources. It would seem that in many areas of Florida agricultural sources would be quite important, and this modeling approach would not seem appropriate for such an important source. Back on page 60 where the model is also described and the model inputs are listed, why is there no information on cropping systems, fertilizer, and manure? The overall modeling approach comes across as quite complex and quite ambitious, while at the same time poorly incorporating agricultural non-point sources. I really wonder how well this modeling can be done for most systems, with enough confidence to make expensive decisions.

In chapter 6 downstream protection values (DPV’s) are discussed with two methods presented to estimate in-stream losses. For N the approach seems reasonable, but for P much is not known about permanent removal of P from streams. The method described seems reasonable, but it requires using the watershed models to get average stream depths on a daily basis. The approach seems quite intensive and takes the modeling results one step further, really making this quite a computational approach, that is still only as good as the overall underlying model. This is despite how many other calculations are added on.

Overall, what comes through clearly to me in the draft document is that we really don’t have that good of data and relationships between N and P and resulting biological effects, so that a wide range of approaches must be utilized, from simple bivariate relationships, to complex simulation modeling. Given that the relationships are not that clear, this suggests that many other factors are involved, and it would seem difficult to account for many of them.

Dr. Robert Diaz

1. General Approach

a) EPA has introduced a general conceptual model in Chapter 2, including the selection of assessment endpoint and indicator variables. What is your perspective of the general conceptual model?

The translation of narrative nutrient criteria into numeric criteria is an appropriate goal for EPA to pursue. Under consideration are numeric criteria for chlorophyll a, total nitrogen, and total phosphorus. Numeric criteria can be precisely defined for these parameters and each identified management unit within a water type can be assessed statistically to determine if nutrient levels and loadings are:

1. At a point for supporting balanced natural populations of aquatic flora and fauna.
2. The degree to which the management unit is out of compliance.

The specific biological components that will be assessed for balance are submerged aquatic vegetation (SAV), benthic invertebrates, plankton, and nekton. Little to no detail is given on what constitutes a balanced natural population. More attention needs to be given to this topic as its definition will drive the numeric values applied to total nitrogen and phosphorus criteria.

EPA proposes three general conceptual approaches:

1. Determining a nutrient reference condition for a system. The use of the term reference implies determining what the nutrient loadings to a system were prior to anthropogenic disturbance.
2. Use of predictive stressor-response relationships and nutrient/algal thresholds. The assumption in following this approach would be that data on nutrient-organism interactions from other regions or countries could be appropriately applied to Florida waters.
3. Numerical water quality models to predict nutrient loadings that would be protective of system biology. The assumption being that a model would be a useful representation of nutrient loadings that will produce a certain water quality.

Detail is given on how these three basic approaches would be applied in each of the Florida water categories. Figure 2-1 is a good flow diagram of what needs to be considered. The first three upper levels (Causal Variable, Response Variable, Water Quality Targets) are deal with at great length, but the bottom two levels (Biological Endpoints, Objective) are not discussed in sufficient detail.

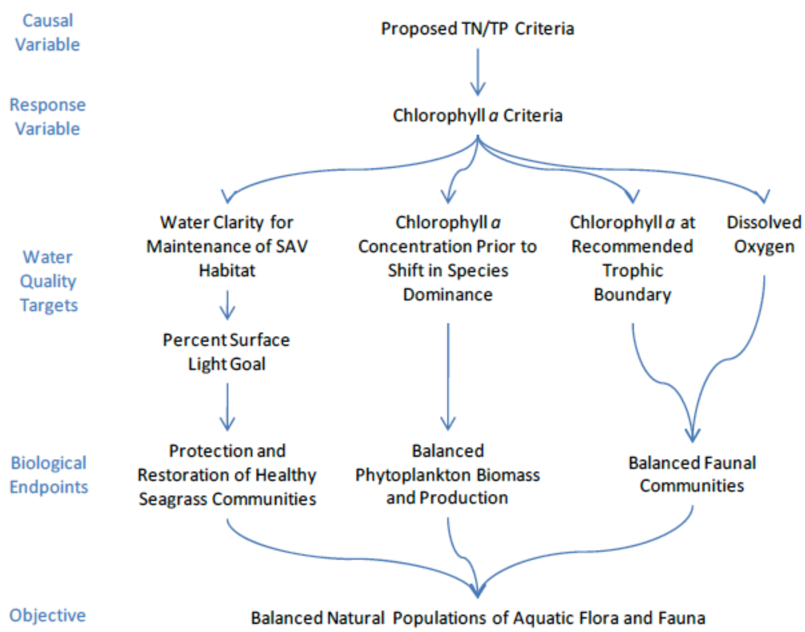

Figure 2-1. Pathways for nutrient effects on estuarine and coastal aquatic life uses.

While we know that nutrients are being delivered to coastal systems far in excess of preindustrial loadings and the negative consequences of these excessive loadings, there is little consideration of the linkage between the top Causal Variable and bottom Objective. How the three general approaches proposed by EPA will incorporate data on what constitutes balanced populations of flora and fauna needs to be expanded. The numeric criteria are being determined to meet the Objective, but there is little detail on the objective. Given that Florida's existing narrative nutrient criterion states:
"In no case shall nutrient concentrations of a body of water be altered so as to cause an imbalance in natural populations of aquatic flora or fauna."
Some consideration for what portion of TN and TP loading in a system is from natural sources and anthropogenic sources is needed. This is particularly important for open coastal waters.

More emphasis needs to be placed on defining what balanced populations are and determining existing conditions of these populations in Florida waters. While we know that reducing nutrients is the key to restoring ecosystems in general, the difficulty lies in setting criteria that can be realistically achieved. In setting TN and TP criteria major consideration needs to be given to all sources of N and P that in combination determine the Biological Endpoints for a system.

b) EPA has delineated the State of Florida into 4 general categories of waters—Florida estuaries, Florida coastal waters, South Florida inland flowing waters, and South Florida marine waters—for purposes of considering approaches to numeric nutrient criteria development . Are these categories appropriate and scientifically defensible? (Note that the details of segmentation of waters within these categories is addressed in subsequent charge questions.)
Separation of estuarine and coastal waters is appropriate given the differences in natural

populations of aquatic flora and fauna between higher salinity coastal systems and lower salinity estuarine and inland systems. Given the complexity of freshwater management the separation of South Florida from the rest of the state is warranted.

Subsections of watersheds, estuaries, and coastal systems for the purpose of setting individual numerical nutrient criteria seems excessive. One problem with this fine-scale approach is that more emphasis may end up being placed on concentration (mg/l) rather than loading (mg/year). It is total loading that flora and fauna populations are responding to. Since supporting balanced natural populations of aquatic flora and fauna is the reason setting nutrient criteria, segmentation of systems should consider spatial distribution of target population. Excessive segmentation also complicates determination of compliance with criteria to be set.

Florida Estuaries (Chapter 3)

2. Estuaries

a) Are the data sources identified appropriate for use in deriving numeric criteria in Florida's estuaries (as discussed in Sections 2.4 and 3.2)? Is the SAB aware of additional available, reliable data that EPA should consider in delineating estuaries or deriving criteria for estuarine waters? Please identify the additional data sources.

Data sources for water quality and system characterization seem adequate. But, the only Biological Endpoint listed in the data sections is SAV. There is no mention of the other Biological Endpoint components in section 2.4. Section 3.2 discusses Biological Endpoints and indicator variables but does not identify data sources.

Appendix B provides a review of the large published literature on the Biological Endpoints. One reference that was not included but should be evaluated, in particular for assessing historical conditions within Florida, is Windsor 1985. It is part of the first national assessment of hypoxia and eutrophication in US waters. And since the EPA nutrient criteria document was released the latest assessment of hypoxia in US waters was published and should be included. Citations are:

Committee on Environment and Natural Resources. 2010. Scientific Assessment of Hypoxia in U.S. Coastal Waters. Interagency Working Group on Harmful Algal Blooms, Hypoxia, and Human Health of the Joint Subcommittee on Ocean Science and Technology. Washington, DC.

Windsor JG, Jr. 1985. Nationwide review of oxygen depletion and eutrophication in estuarine and coastal waters. Final Rpt. to NOAA, NOS, Rockville MD and Brookhaven Nat. Lab. New York. By Florida Inst. Technol., Melbourne.

b) Are the assessment endpoints identified in Sections 2.3 and 3.2 (healthy seagrass communities; balanced phytoplankton biomass and production; and balanced faunal communities) appropriate to translate Florida's narrative nutrient criterion (as cited above) into numeric criteria for Florida's estuaries, given currently available data? Does the SAB suggest modification or addition to these assessment endpoints? A literature review of

endpoints considered can be found in Appendix B.
I find the endpoints to be appropriate, in particular for SAV, but there are other endpoints not mentioned that would be just as good to use. For example, for SAV areal coverage would be important to track. In general, the health and restoration of SAV beds serves as a visible surrogate of total system health. This has been shown to be the case in Tampa Bay where as SAV beds expand other water quality parameters improve, such as hypoxia. Other assessment endpoints are vaguely treated and most are dismissed in Appendix B. Does more detail need to be provided to assess if algae, benthos, plankton, or nekton can work as indicators?

Assessment endpoints that are being considered are:

- Dissolved oxygen
- Algal blooms (is this HAB?)
- Phytoplankton biomass and production
- Benthic and planktonic biological assemblages
- SAV depth distribution

Assessment endpoints that are not being considered are:

- Macroalgae
- Epiphytes
- Macrobenthic and fish indices
- HABs
- Coral
- *Spartina* distribution

It seems that Chapter 2 and Appendix B are at odds over what will and will not be considered. Within each of the assessment endpoints there are many variables that could be measured. More detail is needed on what will be considered. For example, in addition to SAV depth distribution the total areal extend of SAV would be a good metric, as has been shown in Tampa Bay.

c) EPA describes potential approaches in Section 3.3 (reference conditions, stressor response relationships, and water quality simulation models) for deriving numeric criteria in Florida's estuaries. Compare and contrast the ability of each approach to ensure the attainment and maintenance of natural populations of aquatic flora and fauna for different types of estuaries, given currently available data?
Table 3-1 provides a good summary of strengths and weaknesses for the three approaches. The reference condition approach is the most problematic. What would be the baseline for assessing reference conditions? Are there any minimally-impacted water bodies in Florida that fully support Biological Endpoints (Figure 2-1)?

Setting the numerical nutrient criteria based on a percentile cut-point for all water quality data from a Florida system is a good approach. This approach may have the higher probability of success for setting criteria relative to stressor-response relationships, as it would be based on only Florida data. The stressor-response relationships will need to be constructed using data from systems outside of Florida. I would not think there are sufficient studies within Florida to construct the needed regression relationships. However, relative to excess nutrient loadings the response of systems at the national and global level do appear to be similar, which may make the stressor-response relationship good option for Florida waters.

I am not as knowledgeable about water quality models as I should be, but the idea that a model can predict nutrient loads that will be protective of Biological Endpoints is very attractive. Most models are reasonably good at reproducing water quality trends, like dissolved oxygen, but somehow the models need to be linked to assessment endpoints.

Florida Coastal Waters (Chapter 4)

3. Coastal Waters

a) Are the data sources identified in Sections 2.4, 4.1.1 and 4.2 appropriate for use in deriving numeric criteria in Florida's coastal waters? Is the SAB aware of additional available, reliable data that EPA should consider in delineating coastal waters or deriving criteria for coastal waters? Please identify the additional data sources.

I am not as knowledgeable about remote sensing and satellite imagery as I should be, but again the idea that remote sensing can provide data on nutrient and chlorophyll concentrations over large areas through time is very attractive. Is satellite imagery reliable for these purposes so close to the coastline? The bigger question for open coastal waters is; does setting any nutrient criteria make sense? To what extend are coastal nutrients supplied by terrestrial sources and how much of the nutrient flux is related to Gulf of Mexico and Atlantic Ocean sources? If criteria are set how would they be obtained?

b) Is the assessment endpoint identified in Section 4.2 (chlorophyll-a to measure balanced phytoplankton biomass and production) appropriate to translate Florida's narrative nutrient criteria (described above) into numeric criteria for Florida's coastal waters, given currently available data? Does the SAB suggest modification or addition to this assessment endpoint?

I am not sure what balanced phytoplankton biomass and production relative to chlorophyll a is.

c) Does the approach EPA describes in Section 4.2 appropriately apply remote sensing data to ensure attainment and maintenance of balanced natural populations of aquatic flora and fauna in Florida's coastal waters? If not, please provide an alternate methodology utilizing available reliable data and tools, and describe the corresponding advantages and disadvantages.

Without some idea of what balanced natural populations are, I cannot address this question.

South Florida Inland Flowing Waters (Chapter 5)

4. South Florida Inland Flowing Waters

a) Are the data sources identified in Section 2.4 and 5.4 appropriate for use in deriving numeric criteria in South Florida's inland flowing waters (as discussed in Chapters 2 and 5)? Is the SAB aware of additional available, reliable data that EPA should consider in delineating or deriving criteria for South Florida's inland flowing waters? Please identify the additional data sources.

Section 5.4 mentions the use of a macroinvertebrate index, but Appendix B rejected the use of such indices. The index proposed in 5.4 is in a Tetra Tech report that is not available and cannot be assessed.

Because canals are such unusual aquatic systems more needs to be presented on how balanced natural populations are to be assessed. The closest analog to South Florida canals would be in The Netherlands where much of the inland waters flow through canals (locally called ditches). There is some literature for some of the assessment endpoints from Netherland ditches that may be of some use in developing methods for assessing the status of flora and fauna in Florida canals. For example:
Verdonschot, P.F.M. 1987. Aquatic oliogchaetes in ditches. Hydrobiologia 155:283-292.

b) Are the assessment endpoints identified in Section 5.4 (balanced faunal communities, i.e., aquatic macroinvertebrates, and balanced phytoplankton biomass and production) appropriate to translate Florida's narrative nutrient criteria (described above) into numeric criteria for South Florida's inland flowing waters, given currently available data? Does the SAB suggest modification or addition to these assessment endpoints?
It is difficult to determine how minimally impacted canals will be picked or if a canal could at all be classified as minimally impacted. But I think that the use of a combination of a reference percentile cut-point approach and a regression stressor-response relationship approach would be workable.

c) EPA describes two approaches in Section 5.4 (reference conditions and stressor-response relationships) for deriving numeric criteria in South Florida inland flowing waters. Compare and contrast the ability of each approach to ensure attainment and maintenance of balanced natural populations of aquatic flora and fauna in different types of flowing water or geographical areas, given currently available data?
See b). And given the dominance of agriculture in South Florida, why is there only consideration for a 100 m buffer long a canal? I would expect the canal's water quality to be determined by the entire area that drains into it.

South Florida Marine Waters (Chapter 5)
5. South Florida Marine Waters
a) Are the data sources identified in Section 2.4 and 5.5 appropriate for use in deriving numeric criteria in South Florida's marine waters (as discussed in Chapters 2 and 5)? Is the SAB aware of additional available, reliable data that EPA should consider in delineating or deriving criteria for South Florida's marine waters? Please identify the additional data sources.
This area of Florida seems to be data rich.

b) EPA describes two methods in Section 5.6 for using a reference condition approach for deriving numeric criteria in South Florida marine waters (least-disturbed sites or binomial test). Compare and contrast the ability of each approach to ensure attainment and maintenance of balanced natural populations of aquatic flora and fauna in South Florida marine waters, given currently available data?

The two approaches described here are the same as other water types. The binomial test is mentioned for use in each of the four water categories plus the downstream section. This section has the first example. I find that some form of a binomial test would be a good for determining if criteria are being met with some stated precision.

Downstream Protection Values for Florida Estuaries and South Florida Marine Waters (Chapter 6)

6. Downstream Protection Values for Florida Estuaries and South Florida Marine Waters

a. Are the methods EPA is considering for deriving downstream protection values (DPVs) for estuaries (excluding marine waters in South Florida) as described in Section 6.1-6.4 appropriate to ensure attainment and maintenance of downstream water quality standards, given available data? Please describe additional approaches and their advantages and disadvantages that EPA should consider when developing numeric criteria to protect these downstream estuarine waters (excluding marine waters in South Florida), given available data?

Why are atmospheric and point sources excluded from the computation of DPVs?

b. Are the methods that EPA is considering for deriving downstream protection values (DPVs) for marine waters in South Florida as described in Section 6.5 appropriate to ensure attainment and maintenance of downstream water quality standards, given available data? Please describe additional approaches and their advantages and disadvantages that EPA should consider when developing numeric criteria to protect downstream marine waters in South Florida, given available data?

The DPVs need to be linked to criteria derived for estuarine and coastal waters. It is not clear to me how the DPVs for streams and rivers are related to estuarine and coastal protective values. In all it seems a more holistic approach would be needed to assure balanced natural populations. The EPA approach seems to focus on smaller and smaller portions of a system. How will this be managed in terms of numeric criteria for each portion of a system.

Dr. Anne Giblin

Overall I found the document *"Methods and Approaches for Deriving Numeric Criteria for Nitrogen/Phosphorus pollution in Florida Estuaries, Coastal Waters, and Southern Inland Flowing Waters"* to be very comprehensive. The review and synthesis of the literature is extensive and approaches are largely well justified. However, I think that in spite of the difficulties, some additional consideration to protecting health of submerged aquatic vegetation may be needed. There are also places where the document could also be clarified. I have detailed my answers to specific questions below and also included some additional general comments on each chapter.

1) General Approach

a) Assessment of the general conceptual model in Chapter 2

The general conceptual model represents a useful way to represent the relationship between the numeric criteria and the objective of balanced natural populations of aquatic fauna and flora. It is similar to models EPA has used in previous assessments.

I believe classifying water column concentrations of TN and TP as "causal variables", is not quite correct, although I realize this has been done in past assessments. I think it leads to the possibility of some confusion. As the table on page 36 points out, TP and TN ***loading*** are normally considered to be the ultimate driver of ecosystems changes while TP and TN water column concentrations are "associated with influent loading over the long term". Hence I would consider water column concentrations of both TP and TN response variables. The narrative on page 53 also refers to loading as the causal variable and water column concentrations as a response variable. This distinction is important when considering using TP or TN to predict other parameters. Loading data, when available, would be expected to be a better predictor of Chl a, hypoxia and sea grass loss. TN and TP co-vary with Chl a because all are contained in phytoplankton so they are also not completely independent from Chl a. Given the availability of data there may be excellent reasons to use TN and TP concentrations as numeric criteria but they should also be considered response variables. In addition, it would also be useful to characterize these variables in a little more detail. I realized the specifics are yet to be determined but would these numbers be calculated on surface values, depth integrated samples, discrete depths?

The last sentence on page 32 strongly suggests to me that numeric criteria for TP, TN and chlorophyll a will be the only ones established. Yet, there is a numeric dissolved oxygen standard in Florida. I'm not sure why this isn't just used as an additional numeric criteria. Instead, EPA plans to address this standard using "the relationship between DO and TN or TP". DO is a critical variable, is frequently measured, and is calculated in all water quality models. It is also strongly affected by physical processes so one will not find general relationships between DO and TN, or TP across systems. To add to the confusion in chapter 3 it does appear that oxygen will be used as a criteria in some cases. I would suggest that this section be clarified.

On page 38 there is a discussion of changes in phytoplankton and zooplankton assemblages which ends with "EPA is considering using Chl a as an indicator of the changes in phytoplankton species composition". I believe I see the intent but didn't find much support for this idea in the paragraph. Overall, Chl a is a fairly insensitive way to look at composition changes. Perhaps a better way to deal with this is point out that by protecting against biomass increases the species composition if less likely to change than if there were large changes in biomass.

In this section sea grass beds are being protected by maintaining water clarity by limiting the accumulation of Chl a in the water column. According to the Appendix, numeric criteria for light levels were not proposed because they can be affected by non-nutrient related changes. (although they are then considered in Chapter 3 see below). No numeric criteria directly related to sea grass, macroalgae, or epiphytes is being proposed. I am concerned that the Chl a variable will not be responsive enough to protect sea grasses. In systems where the nutrients are largely taken up by the phytoplankton this will work. But there are systems where even with nutrient increases the water column Chl a levels stay at fairly low levels due to short residence times, but macroalgae proliferate. Hauxwell et al. (2001, 2003) found that light levels in macroalgal mats prevented young eel grass shoots from being established. Epiphytes can also increase in systems where Chl a levels may still be fairly low. EPA could consider an approach such as was recently taken in New England (Latimer and Rego 2010 Estuarine, Coastal and Shelf Science). Data on eel grass loss was obtained for a number of estuaries and nutrient loading calculated. This study found eel grass loss began to occur at N loads above 50 Kg ha^{-1} y^{-1} and eel grass disappeared at 100 Kg ha^{-1} y^{-1}. It may be possible to develop a similar relationship for Florida sea grass systems.

b) *Are the four general categories appropriate and scientifically defensible*

The classification scheme is appropriate. Florida's inland waters are highly managed and present challenges. The approach taken here is a good way to classify the system.

2) Florida Estuaries (Chapter 3)

a) *Are the data sources appropriate? Other possible data?*
To the best of my knowledge they have obtained good data sets.

b) *Are the assessments endpoints appropriate to translate into numeric criteria given available data? Should there be an addition or modification to these endpoints.*

Some aspects of these endpoints contradict some of the statements in Chapter 2. As discussed above, the issue of sea grasses needs careful thought if protection is to be achieved. This section describes how depth to colonization may be a good way to indicate sea grass condition. This is one of several metrics that could be useful. If there is historical data on sea grass depth distributions than changes in nutrient loading that have occurred could be used to help develop loading criteria in the same way that Latimer and Rego did using areal extent.

As mentioned above, I think Chl a will provide important information on biomass changes but will be less useful for species changes. Oxygen will be a useful criteria for balanced faunal communities and should be included.

c) *Compare and contrast approaches for deriving numeric criteria.*

EPA has clearly describes several possible ways in which the nutrient criteria can be derived. All have strengths and a combination of approaches are proposed here. Because there are relatively few systems EPA is proposing a specific criteria for each estuary. This is ambitious but an excellent approach as the estuaries are quite different and this will allow for the best possible criteria to be developed.

The first approach (reference condition approach) is to define the unimpacted reference condition either base upon historical information or by using data from unimpacted systems. EPA has determined it has not identified impaired estuaries for which historical data exists to create a suitable reference condition. This may be the case for the water column parameters but it does appear that historical data is available for sea grasses (Fig 3-6). This may be a matter of semantics but using this historical data on sea grass extent to help set the numeric criteria seems to be very useful. Using this information will obviously be a two step process since the loads present historically will have to be calculated (water quality simulation models). This could be compared to data achieved using the reference condition approach derived by using existing water quality data alone.

Certainly an analysis of the TN, TP and Chl a values in unimpacted systems will be very valuable and should be done as soon as possible. The relationships and the variability in this data will be informative and may suggest which of the methods of calculating the data (geometric concentrations or average and upper 25%) to develop the relationship makes the most sense.

A second method is a stressor response method. Regression models have long been an excellent way for scientists to uncover relationships but often have a great deal of scatter. In this case, applying a single regression model to the range of systems does not seem warranted when more powerful approaches are available. I agree with EPA (page 53) that models which span a large range of systems, such as that put forth by Steward and Lowe (2010) will probably not provide estimates of protective loads with sufficient precision.

When applied on the basis of an individual estuary, stressor response can be very useful for developing specific criteria. They often are also useful in determining if water quality models are behaving properly. The Janicki and Wade (1996) model demonstrates the power in this approach. EPA should continue to pursue these types of approaches when ever practical.

Water quality models have been used with great success in some estuaries and open water systems, but there have also been failures. To actually set numeric criteria in these systems getting the hydrodynamics correct is critical. I am not knowledgeable enough to evaluate the hydrodynamic models proposed here. I do know for shallow systems the finite volume approach taken by Chen's group at U. Mass Dartmouth has proved extremely successful and computationally efficient. However, there appear to already be models available for a number of these systems and these may be working well.

The water quality model chosen here (WASP) has been widely used. The choice of parameters will be important. Phosphorus uptake and release from the sediments is difficult to model and unless done well, could lead to poor results. The document (page 64) points out that the benthic fluxes require a great deal of data. They suggest that benthic fluxes could be specified as boundary conditions but this will lead to poor results. These fluxes are quite responsive to changes in loading. In addition, the sediment responds to changes in loading by altering P adsorption coefficients and denitrification rates which feedback into the water column values. Finally, these models don't include sea grasses and only a rudimentary job with macro-algae so they may be of limited use in the very shallow and pristine systems. However, there is a great deal of existing water quality data which should allow for calibration of the model. An important feature to examine would be whether or not the calibrations between systems require very different parameters in each system to be calibrated. This would suggest that some key processes are missing and that the model may perform well for hindcasting or forecasting.

I am not familiar with this watershed model but a good watershed model is critical, not only to this effort but for setting TMDLs later on. For many of these systems it appears that good data on discharge and loading are available to parameterize the model.

3) Florida Coastal Waters (Chapter 4)

a) *Are the data sources appropriate?*

I am not familiar enough with this type of data to know.

b) *Is the assessment of endpoint appropriate given currently available data?*

I don't feel I have the background knowledge on this system to comment at this in detail at this point. There appears to be somewhat higher Chl a values off of estuaries which may reflect increased nutrients over historical values. A key question of course is has this always been the case? There is also no other link to the Chl a values and other resource values (benthic communities, sea grass) to know at what point the increase in Chl a is affecting the resource.

c) *Does the approach appropriately apply remotely sensed data to assure goals?*

EPA states it has not identified an "impaired" water body where historical data is available but there is a great deal of earlier remotely sensed data. Clearly the input of nutrients from land can affect coastal waters but it would very valuable to know if the areas which have seen an increase in the discharge in land based nutrients have seen a change in chlorophyll over those that have not.

The goal is to protect the coastal waters. It would seem that values that are protective of the estuaries should largely be protective of off shore water.

4) South Florida Inland Flowing Waters (Chapter 5)

a) *Are the data sources appropriate, are there others?*

The Florida Coastal Everglades has a considerable amount of data , much is marine but there is some freshwater data.

b) *Are the assessment endpoints appropriate ?*

For the freshwater sites, given the lack of data TN and TP in the least disturbed sites is a good first step. The faunal assessment would be helpful but sites surrounded by wetlands may differ in ways not simply related to nutrients. Given the high degree of modification of these channels EPA has taken a logical approach though. I agree that there is limited data to know what a "protective" level of Chl a would be for these systems. This approach also ignores macroalgae and vascular plants- aren't these frequently a water quality problem in these systems?

c) *Compare and contrast the approaches for deriving numeric criteria.*

My comments above, and on the general approach would apply here as well.

5) South Florida Marine waters (Chapter 5)

a) *Are the data sources appropriate and are there others?*

The Florida Coastal Everglades project has a great deal of data, on sea grasses, nutrients, primary production, etc. see http://fcelter.fiu.edu/data/FCE/

b) *Compare and contrast approached for deriving numeric criteria*

I am not a statistician but the use of the binomial test to maintain current conditions seems more appropriate for a variable like Chl a. As pointed out in the report, this metric addresses the issue of infrequent blooms as well as gradual changes in the mean.

6) Downstream Protection Values (Chapter 6)

a) *Are the methods appropriate for estuaries?*

The general approach is appropriate although I strongly urge it be base upon loads not concentrations (both were suggested on page 122). Changes in climate, water use, etc. can all impact water delivery and change loading in complex ways. Determining the loading on a segment basis will be very important if protective criteria are to be achieved. However, a number of parts of this description were confusing. For example, I'm not sure why the protective load would exclude loads from direct atmospheric deposition or point sources?

The in-stream loss rates are critical factors that are very difficult to estimate as detailed here. However, there has been considerable progress on empirically estimating losses through experiments and modeling. Before using the simplified approach (a single coefficient for N for example) it would be worthwhile attempting to determine some site specific coefficients for the region. The large scatter in the values from some of the national studies reflect differences in geology, etc. that are greater than would be in this area alone. Similarly, P storage needs to be carefully considered.

b) *Are the methods appropriate for downstream water quality for marine waters in South Florida?*

Some of the same issues brought up above apply here. First, unless the criteria is set based upon loads, rather than concentrations, the marine systems may see increased loading with changes in climate, etc. However, I realize the difficulty of the implementation so perhaps concentrations could be established which would be provisional on flows within a certain range.

Dr. Judith Meyer

1. General Approach (Chapter 2)

a) What is your perspective of the general conceptual model?

(1) I have a fundamental concern with the general approach described at the top of p. 30 and implied in the rest of the document. The guidance documents describe 3 general approaches (reference conditions, stressor-response, and modeling) and note that each is appropriate but with different data requirements. My concern is that not only does each have different data requirements but also each has different assumptions, limitations, and uncertainties. Greater confidence in the results seems likely if more than one method is used and guidance is provided as to how to combine the results of different analyses. Using only one approach when all have limitations and uncertainties is not making optimal use of all the information and analytical capabilities available. Yet in the document, it appears that use of only one approach is acceptable.

(2) Section 2.1.2, numbers 4 and 6: Some earlier clarification of terms as used in this document is needed. Specifically how are the following different: "causal and response variables", "assessment endpoints", "measurement endpoints", "indicators". This is done to some extent in section 2.3, but I think that discussion should be moved earlier in the document as it is fundamental to understanding what is being proposed. There is also discussion of the "concepts of assessment endpoints and indicator variables" in Chapter 3. That belongs in Chapter 2.

(3) Section 2.1.2, number 7: It is unclear whether BOTH reference and modeling approaches are recommended or if only one will suffice. I think it should be the former as noted under (1) above. Where are stressor-response relationships in this description? It is also unclear what variables are to be analyzed in determining the reference condition – causal and response variables? assessment endpoints?

(4) The explanation in the paragraph above Figure 2-1 implies that low oxygen is the only mechanism of nutrient impacts, which is obviously not the case.

(5) Figure 2-1 is not showing pathways for nutrient effects. It is showing the intended consequences of setting regulatory criteria. Clearly it all hinges on chlorophyll, which I presume is water column chlorophyll, although that is never stated. It is not clear what is meant by "balanced phytoplankton biomass and production" or "balanced faunal communities". I recognize that this reflects the current narrative criteria, but an explanation of what EPA thinks this means is important.

(6) Table 2-1 is an excellent summary of what is included in Appendix B, which I found to be a thorough and thoughtful review of the literature. However nowhere in that table does one find "balanced phytoplankton biomass and production" or "balanced faunal communities", which are then proposed as two of the assessment endpoints in 2.3.2.

(7) 2.3.3.4 Appendix B provided an excellent review of the relevant literature. Where I found it lacking was in how that information was used to accept or reject certain indicators for nutrient effects, which is what is referred to here. Specifically, I found the rationale for using or not using the following indicators to be weak: clarity (B-5; the same criticisms apply to the chlorophyll-a measure proposed for seagrasses); the phytoplankton endpoints (B.3.4) are not adequately supported; the reasons for rejecting an indicator based on macroalgae (B.4.4) are inadequate. The discussion of epiphyte response (p. B-24) asserts that impacts to that assemblage occurs before impacts to seagrasses and could serve as an early warning of nutrient impacts. This argues against use of the seagrass-based endpoint.

(8) 2.3.2.1 This assessment endpoint is not clearly explained. Water clarity requirements (water column chlorophyll a?) are to be computed as % of surface irradiance required at a selected depth. What %? How is depth selected? How is water column turbidity from suspended sediments or colored DOM factored in? Clearly excess epiphytic growth is another nutrient-related stress on seagrasses; this approach seems to assume that water column chlorophyll and epiphytes will respond similarly to nutrients.

(9) 2.3.2.2 I question the rationale for using chlorophyll a as an indicator of changes in phytoplankton species composition. This requires considerably more justification than is presented in this document.

(10) p. 42: Does the National Elevation Dataset have adequate resolution for a state as flat as Florida?

(11) Nowhere in this document is N:P ratio considered. Has this been deemed irrelevant for these ecosystems?

b) Is the division into 4 categories (estuaries, coastal, South Florida flowing, South Florida marine) scientifically defensible?

(1) It is defensible, but has not really been defended in this document. Just a paragraph explaining the rationale for the 4 categories is needed.

(2) The more detailed classification described in Appendix D was produced in May 2010. The report implied that more work was forthcoming. What has been determined in the past 6-7 months?

2. Estuaries (Chapter 3)
a) Are data sources appropriate for estuary delineation and criteria development? Other data sources that should be considered?

(1) One strength of the proposed approach is that system-specific and even estuary sub-segment endpoints will be determined.

(2) The data sources seem appropriate (but note my above concern with the National Elevation Dataset). To what extent do groundwater inputs correspond to the watershed boundaries that are identified? I do not know what data are available for groundwater flows in the state, but I find the absence of any discussion of this to be troubling. Similar concerns arise in the discussion of watershed models in 3.3.3.1.

(3) The map in Figure 3-2 is confusing because it combines both estuaries and coastal systems and I guess also South Florida marine. It would be clearer if separate maps were presented for each.

b) Are assessment endpoints (healthy seagrass communities; balanced phytoplankton biomass and production; and balanced faunal communities) appropriate? Suggested modifications?

(1) I am really confused as to what is proposed as the assessment endpoint for seagrasses. Reading the description on p. 49 makes it sound as though the depth of seagrass colonization will be measured (i.e. how deep are they observed), but that is not what was described in Chapter 2 [see comment 1.a.(8)]. Then I read in Section 3.4 that depth will be determined from available data on occurrence laid over bathymetry. It is not clear how this will result in nutrient criteria, and no mention of this is included in the Pensacola Bay example.

(2) Although "healthy" seagrass communities have not been defined, I would think that an absence of invasive species would be one aspect of "health." Yet I have read nothing about invasive species. Is that not an issue with seagrasses, or not one that is indicative of excess nutrients?

(3) "balanced phytoplankton biomass and production" The use of chlorophyll as an indicator of changes in phytoplankton species composition has not been defended. Furthermore, how is a measure of water column (?) chorophyll going to indicate something about the "proliferation of macroalgae, epiphytes, or nuisance algal species" as implied on p. 49?

c) Three approaches (reference conditions, stressor response relationships, and water quality simulation models) are described. Compare and contrast the ability of each approach to ensure the attainment and maintenance of natural populations of aquatic flora and fauna for different types of estuaries, given currently available data.

(1) As noted above [comment 1.a.(1)], the three approaches (reference condition, stressor-response, simulation modeling) should not be used independently; they should be used in combination to develop criteria. Table 3-1 is a good summary of when each is applicable, and when more than one approach is applicable, more than one approach should be used.

(2) It sound as though the reference condition approach will require that water quality conditions remain as they are at present in unimpaired estuaries. Yet the estuary may be able to withstand higher nutrient inputs and concentrations without loss of "balanced flora and fauna". This approach would seem to set a lower bound for nutrient criteria and is a further argument for use of more than one approach to establish nutrient criteria.

(3) Are the number of samples and years of data adequate for all of the estuaries listed in Table 3-2? Some indication of length of time and number of samples considered adequate would be useful – or at least flag those estuaries where there does not appear to be adequate number of samples. I realize that in part depends on the nature of the data (e.g., variability), but presumably that information is available for the estuaries listed in that table.

(4) One concern with all of the approaches is that I see little discussion of seasonal differences and of interannual variability. Estuarine vulnerability to added nutrients probably varies by season; yet this does not seem to be taken into account in this discussion. Some consideration of extent to which the period of record or period for which models will be run (e.g., LSPC from 1995-2009, p. 60) and evaluated (e.g., LSPC, p. 76) includes a range of hydrologic conditions would seem to be needed in assessing data and model adequacy for the analysis proposed.

(5) Indicators of ecosystem function are proposed as part of model evaluation (p. 65). Will model output be compared with data? Are those data (e.g. primary and secondary productivity) available?

3. Coastal Waters (Chapter 4)

a) Are the data sources identified appropriate? Is the SAB aware of additional available, reliable data?

(1) Remote sensing data will not be used for coastal waters from Apalachicola Bay to Suwanee River (p. 77) because of interference from colored DOM, and criteria development in these waters is supposedly covered in Chapter 3. This sounds reasonable, but I did not find that information in Chapter 3.

b) Is the assessment endpoint - chlorophyll-a to measure balanced phytoplankton biomass and production - appropriate given currently available data? Does the SAB suggest modification or addition?

(1) Nutrient criteria are not going to be developed for coastal waters (p. 78). Water column (?) chlorophyll-a in unimpaired areas (reference condition approach) will be used to assess if coastal areas are maintaining water quality. Chlorophyll-a is an imperfect measure of phytoplankton biomass, which this document acknowledges. Yet remotely sensed chlorophyll-a seems to be a reasonable and logistically feasible endpoint for the extensive areas needing to be assessed. However, it cannot be used as an indicator of species composition, which is implied in this document. The scientific basis for this statement is not convincing.

c) Does the approach appropriately apply remote sensing data? If not, please provide an alternate methodology and describe the corresponding advantages and disadvantages.

(1) Does the past 10 years represent a long enough time period for the reference condition approach? Is this period really an appropriate baseline if nutrient criteria have not been in effect? It seems that this approach would keep conditions from getting worse but not improve things.

4. South Florida Inland Flowing Waters (Chapter 5)

a) Are the data sources identified appropriate? Is the SAB aware of additional available, reliable data sources?

(1) The proposed inventory of inland flowing waters that catalogues and distinguishes natural streams and canals should provide very useful information.

(2) The document states that criteria for Class IV waters will not be determined (p. 98). However, establishing DPVs could clearly result in criteria for Class IV waters.

b) Are the assessment endpoints -- balanced faunal communities, i.e., aquatic macroinvertebrates, and balanced phytoplankton biomass and production -- appropriate? Does the SAB suggest modification or addition?

(1) Use of the multimetric macroinvertebrate index is the most direct assessment endpoint for balanced faunal communities that has been proposed in the entire document. It is a much more direct assessment of this than is the DO measure proposed for estuaries.

(2) I do not understand the reference to “limited evidence” for establishing criteria for chlorophyll-a in canals. Has chlorophyll-a not been measured there? Is perhaps benthic chlorophyll-a not a more appropriate measure in these systems? I find inadequate information in this document to evaluate the statement on “limited evidence.”

(3) The 100m buffers proposed for use with the LDI (pp. 105-106) may be too limited, particularly where stormwater pipes convey runoff from distances much further than 100m.

c) EPA describes two approaches -- reference conditions and stressor-response relationships. Compare and contrast the ability of each approach in different types of flowing water or geographical areas, given currently available data.

(1) I have some concern with use of annual average concentrations (p 107). Seasonal differences in response to nutrients can be expected so a nutrient criterion could be different in different seasons. I have not seen a discussion of this in this document. Furthermore, as mentioned in this section, duration of conditions of elevated nutrients is not captured by the use of annual average concentrations that can be exceeded in consecutive years.

(2) The results of a distribution approach (p. 108) is sensitive to the distribution of sites along the disturbance gradient. If a larger proportion of the samples are from more disturbed sites, then using the lower percentile to set the criteria will result in a higher number than if a larger proportion of the samples are from less disturbed sites. Some requirements with respect to the distribution of sites along the disturbance gradient are needed.

(3) Some of the variability in the stressor-response relationship could be a result of season. This should be investigated, and it may lead to the formulation of different criteria in different seasons.

5. South Florida Marine Waters (Chapter 5)
a) Are the data sources identified appropriate? Is the SAB aware of additional available, reliable data?

(1) Are seagrass criteria not being proposed for this ecosystem type? Chlorophyll-a based criteria are derived using a reference condition approach, but there has been no mention of relating this to seagrasses.

b) EPA describes two methods for using a reference condition approach -- least-disturbed sites or binomial test. Compare and contrast the ability of each approach, given currently available data.

(1) I do not think it should be either/or. Both methods should be used with a mechanism for how to interpret any difference in the results.

6. Downstream Protection Values for Florida Estuaries and South Florida Marine Waters (Chapter 6)
a) Are the methods for estuaries appropriate given available data? Please describe additional approaches and their advantages and disadvantages that EPA should consider.

(1) Inclusion of DPVs is an important component of this document. This represents progress in developing criteria that are protective of water quality.

(2) When assessing fraction delivered, some consideration of seasonal variation is essential. The fraction delivered will vary greatly depending on whether it is low flow conditions where instream removal processes can remove a great deal vs. intense storms where those instream removal capabilities are overwhelmed. There is little scientific justification for using a constant loss rate.

(3) Figure 6-1 does not make it clear that estimated protective loads coming from tributaries are reduced based on direct point source and atmospheric inputs to the estuary. Also, how this load is apportioned among tributaries is not included in this figure.

b. Are the methods for marine waters in South Florida appropriate given available data? Please describe additional approaches and their advantages and disadvantages that EPA should consider.

My responses to “6a” apply here as well.

Editorial comments
Executive summary: Where is urban stormwater in the first paragraph?
Figure 1-1: No human health impacts are included; e.g., nitrate and algal toxins in potable water supplies.

Dr. Mark R. Noll

The EPA presents proposed methods and approaches for determining TN and TP concentrations and loads in the estuarine and coastal waters of Florida and the inland flowing waters in South Florida. The approaches presented in the draft document correctly take into account the maintenance of whole ecosystems, and further recognize the internal variability of seemingly similar systems. In determining TN and TP values that are protective, the authors suggest a segmented approach to the analysis of load distribution of contributing watersheds. This approach is mindful of the many factors, both natural and anthropogenic, that influence the transport of N and P through a watershed to potentially fragile receiving waters. The segmented approach recognizes the practical need to apportion contributions from different contributing areas of a watershed that may relate to additional factors such as land use. Overall, the methods and approaches presented provide a fundamentally sound methodology for the management and protection of the environment.

The methods and approaches presented, however, need additional explanation that will make the document more instructive to future users. The use of several models for the determination of Downstream Protection Values (DPVs) is a valid approach, but additional guidance will be required to enable end users to develop predictive models that are sufficiently protective given the inherent variances of model outputs. Furthermore, the draft document does not seem to emphasize the use of geographic information systems (GIS) in the evaluation and modeling of watersheds. The use of GIS may be implied given the models being used, but the value of this tool should be emphasized. While the use of predictive models is an important component to ecosystem management programs, it is crucial that oversimplification of input parameters is avoided. The draft document presents some simplified approaches and assumptions. For example, the concept that P may be permanently retained within a watershed may be overly optimistic. While the limitations of the draft document are important, they are easily corrected, and do not severely limit the usefulness of the methods and approaches presented. Rather, they are precautionary in nature so that future users may apply them most effectively.

Dr. Hans Paerl

1. General Approach:

The approach is a well-focused, comprehensive and thoroughly-documented effort at developing and applying numeric criteria aimed at maintaining and (in the case of degraded waters) improving water quality and habitat conditions in Florida's (southern) inland, estuarine and coastal waters. Human encroachment continues to intensify and impact these waters. The issues and problems facing these waters are clearly articulated and systematically addressed. I found the rationales for taking on this task well-laid out and justified. There were very few stones left unturned in terms of identifying and addressing key drivers/stressors impacting water quality, habitat condition and aquatic resources. A wide range of trophic levels and states are addressed in terms of responses and their interactions (e.g. planktonic vs. benthic primary production impacts and interactions, physical-chemical-biotic interactions, including the roles of turbidity, flow, temperature). The classification scheme and waterbody delineation criteria used to evaluate criteria for impacts seemed reasonable and defendable. Response indicator variables to be used as criteria of impacts are logical, justifiable (from individual system and process perspectives) and, in most cases, interpretable and quantifiable. Data collection and management approaches will allow for both system-level and comparative (regional) approaches and facilitate multi-media (field evaluation, remote sensing, modeling) analyses and evaluations. Overall, the approach is logical, clearly articulated and adequately justified, based on the issues and resources to be addressed.

There is the issue of whether to use TN/TP or "reactive N and P" (i.e. DIN and DIP) as forms of nutrient enrichment and linking these to specific effects (i.e. primary production, biomass as chlorophyll *a*, and cascading effects such as food web alterations and hypoxia). This issue has been the subject of considerable research, discussion and controversy for numerous decades. Much of the uncertainty regarding whether to use TN/TP or more "reactive" dissolved forms of these nutrients revolves around the bioreactivity and roles organic forms of these nutrients play. The degree(s) of bioreactivity are the current subject of research. Bioreactivity may be system- (or even components of systems) specific, adding to the complexity (and uncertainty) of measuring responses and their impacts on water quality and habitat condition. It will be important for the Panel to discuss this issue in the context of developing numeric nutrient criteria for nutrient-sensitive waters, both in Florida and nationally.

2. Estuaries: The issues facing estuaries in the context of nutrient enrichment and its ecological and biogeochemical impacts are clearly spelled out and addressed. While it is difficult to characterize and delineate estuaries (as well as within-estuarine areas) along geographic gradients for criteria development, endpoint assessment and use of specific indicator variables, the approach taken here seems reasonable and, to a large extent defensible. Estuaries are notoriously-difficult to delineate or "pidgeon-hole" for indicator and response variables, in large part because highly individualistic hydrology (freshwater inflows), morphology, tidal exchange and resultant water residence times or "water ages" strongly influence the systems' biotic and habitat responses to nutrient inputs. EPA has recommended a delineation approach based on "natural geographic limits of estuarine basins and their associated watersheds". This seems reasonable form a land use-nutrient input-primary produce r response perspective. However, even within a single estuarine system, there is tremendous spatio-temporal variability with regard to the amount and type of primary production and plant community structural response. An example in Florida is the complex St. John's River System in Central and Northeastern Florida, which exhibits a range of water quality and habitat degradation problems, rnaging from upstream toxic cyanobacterial blooms to downstream dinoflagellate blooms and periodic hypoxia.

When incorporating flow and water residence time, the most direct response to N and P enrichment, namely chlorophyll *a* (phytoplankton and periphyton) biomass, can now be reasonably-well predicted based on knowledge of limiting nutrients and flow/residence time regimes for specific segments of an estuary like the St. Johns. The complexity of system-level responses becomes more challenging when examining these responses on the habitat-level to less directly (linearly)-related indicators and criteria (e.g. hypoxia), as these are dependent not only on flow but also vertical stratification, inputs of substances other than N and P (i.e. organic matter). Lastly, Florida estuaries exhibit a strong interplay between planktonic primary production responses to nutrient inputs (including algal blooms) and benthic production and community structure responses (e.g. seagrasses, benthic microalgae), due to the complex interplay of nutrient enrichment, changes in water column residence time and transparency. The authors of this chapter recognize these challenges however.

Whether or not to use TN/TP or DIN/DIP as the nutrient loading and concentration variable continues to be a challenge in estuarine ecosystems. The roles of organic N and P remain largely unknown in terms of plant activity and community structural responses. However, it is believed that DON and PON are a least in part biologically reactive in estuarine ecosystems. Given that residence times and hence the period available for uptake, degradation and cycling of N and P compounds, can vary dramatically in Florida estuarine systems, it is probably most inclusive to use TN and TP values in order to examine the potential extent of bioreactivity and trophic/biogeochemical responses to these nutrients. Also, from a watershed and airshed loading perspective, it is prudent and to account for *all* forms of N and P being derived from these sources.

In this chapter, the authors have thoroughly considered and incorporated the confounding effects of hydrology, light (turbidity), trophic (grazing), biogeochemical (internal cycling) of these nutrients in developing numeric criteria for evaluating ecosystem-level responses. Up to date modeling efforts have gone into considering and evaluating these complex responses in estuarine systems. This chapter also does and excellent job of incorporating and linking land use (and changes therein) in assessment of estuarine responses to nutrient loads under hydrologically-variable conditions. State variables, relevant processes, boundary conditions and indicators to be used in the applications of models are adequately addressed and included. The current and future use of remote sensing is also addressed in the context of characterizing and quantifying drivers and response indicators.

One aspect of hydrology that warrants additional attention is future water freshwater withdrawal scenarios, and their potential impacts on nutrient loads, concentrations and downstream estuarine hydrologic, water quality and habitat conditions. There is little doubt that Florida will face increasing freshwater demands as urban centers, agriculture and some industries continue to grow. What will the effects of increasing water withdrawals be on nutrient-water quality and habitat interactions/relationships? This looming issue will need to be addressed in the context of estuarine responses to nutrient inputs in the near and distant future.

3. Coastal Waters: Florida contains highly diverse, resourceful coastal waters (defined in this document as "marine waters up to three nautical miles from shore, but excluding estuaries……". Assessing effects and ecosystem-level responses to N and P inputs is a challenge in these waters, in part because they are trophically and biogeochemically diverse, and also because (with few exceptions) they are not routinely monitored for water quality using indicators prescribed in this document (e.g. chlorophyll *a*, turbidity, dissolved oxygen, etc.). To address this challenge, the EPA is considering a "reference-based approach with satellite remote sensing chlorophyll *a* observations to derive numeric values that translate Florida's narrative criteria and ensure support of a natural balanced population of aquatic flora and fauna. This

approach is likely to be quite effective in Florida coastal waters, because they are optically amenable to remote sensing of chlorophyll, color (CDOM) and turbidity. Remote sensing technology has evolved sufficiently to begin using calibrated imagery for estimating chlorophyll. Furthermore, this technology has been used successfully to detect, quantify and characterize (as to potential linkages to nutrient supplies) harmful algal booms in these waters as well as regions beyond 3 miles offshore.

The proposed establishment and use of reference conditions is clearly spelled out and justified in this chapter, although it isn't clear how much useful; "historic data" there are against which to compare current conditions and changes. This can be clarified however. Incorporating the stressor and response data into coastal water quantity and habitat response models is another challenge. Overall, the approach to be used is well laid-out and rationale/potential criteria and products clearly articulated.

4. South Florida Inland Flowing Waters:

South Florida's inland flowing waters have a long history of being highly manipulated and managed, and in this regard represent a special challenge to developing numeric criteria for nutrient pollution impacts. The classification of these waters is highly anthropogenically-influenced; largely a product of terrestrial and hydrologic modifications. In this regard, there is logic to classifying these waters according to basin and sub-basin soil types and land uses. An additional challenge is incorporating groundwater hydrologic/nutrient dynamics, which have also been altered, but are likely to be very important in determining nutrient pollution sources and impacts. The proposed classifications in this chapter appear reasonable as they incorporate surface and subsurface flow regimes and flow lines as well as soil types and human agricultural and urban impacts (i.e. land use). A substantial amount of effort will be put into identifying and quantifying stressor-response relationships in these waters using correlative/regression analysis. Considering the difficulty of working across the surface-subsurface interfaces in deriving nutrient loading estimates, as well as effects of these loads, the authors have done a good job of addressing these challenges. This section could however benefit from closer process/response connections (including applying modeling approaches) to receiving estuarine and coastal waters.

5. South Florida Marine Waters:

South Florida marine waters (Florida Bay-Keys region) represent a formidable challenge (as opposed to other FL coastal waters) because the freshwater inflows and nutrient inputs to them are highly manipulated and managed. EPA is considering using the Southeastern Environmental Research Center (SERC) water quality network, which has an extensive data base on TN, TP, and chlorophyll *a* as a primary data source for evaluating nutrient input/pollution numeric criteria. This seems reasonable and logical, as it appears that the SECR data set is the most extensive and comprehensive for the region. Water quality effects of nutrient inputs and well as hydrologic alterations tend to be regionalized in sub-basins of Florida Bay, which are semi-compartmentalized (flow and residence time-wise) and hence show individual different sensitivities to nutrient inputs. Nutrient limitation also seems to vary among various regions and sub-basins of Florida Bay. Nutrient limitation and enrichment are closely controlled by the input of P via exchange with the Gulf of Mexico to the west and N and P inputs from freshwater pollution from the Everglades, groundwater and channel inputs to the north and east. Therefore, there is a great deal of regionalization in terms of nutrient sensitivity and enrichment, presenting a challenge to evaluating nutrient pollution criteria.

Prior work in these waters has shown that statistical modeling analysis and hierarchical cluster analysis were effective in delineating these regions with respect to nutrient sensitivity and responses to nutrient enrichment. Furthermore, these types of analyses were able to separate benthis SAV habitats into "ecologically distinct regions". Discriminant function analysis coupled to monitoring was able to predict benthic habitats and their responses to nutrient enrichment scenarios. This provides some promise that regional criteria of nutrient sensitivity may be developed in these waters. The entire approach needs to be better developed however, and may require additional monitoring end experimental work (nutrient limitation and sensitivity assays) to evaluate and verify.

EPA is considering two approaches for deriving criteria using a reference condition approach: 1) Statistical distribution of water quality indicator variable data and 2) statistical distribution of raw data and evaluation using a binomial test. This would potentially enable EPA to determine the average concentration and an upper percentile concentration. These seem like reasonable approaches and should be encouraged and tested. The proposed binomial tests should be evaluated for its ability to provide information on distribution of data during the observation and assessment period, whether the approach is "resistant to influence of extremely high values", and if the approach is affected by censored or otherwise omitted data. The approach seems reasonable and potentially quite useful in establishing and testing criteria.

6. Downstream Protection Values:

The Downstream Protection Values (DVPs') that EPA will attempt to develop, based on numeric criteria for streams in Florida in order to protect the estuarine water bodies impacted by N and P pollution from the watershed, appear to be a logical product of the overall effort to develop numeric nutrient input criteria for estuarine ecosystems. The DVPs will be based on limits on TN and TP loading rates that are required in order to ensure "balanced natural populations of aquatic flora and fauna in estuarine waters". This in essence follows the procedures and analysis proposed in chapter 3. The computational steps needed t achieve the DVP appear essential, sound and logical, although this is not in the area of expertise of this panel member to thoroughly evaluate. The inclusion of flow corrected rates of nutrient loading (and nutrient retention and attenuation) as well as estuarine water quality responses, is highly appropriate and essential in this highly hydrologically-variable region. It is clear to this panel member how the DVPs may be calculated based on a fairly direct response indicator like chlorophyll *a*, but not necessarily to a less direct (and often non-linear) response indicator like hypoxia and associated habitat degradation. Lastly, it isn't clear how the DVP will deal with reductions of one nutrient (e.g. P in upstream P limited freshwater systems) vs. dual nutrient input reductions, which are often required along the freshwater to marine (riverine-estuarine-coastal) continuum.

Dr. Kenneth Reckhow:

- Overall, I find the defensible scientific linkage between candidate criteria/levels and designated use to be weak. Developing numeric criteria is a good idea, but these criteria need to be predictive of designated use, and I don't find that to be scientifically supported (particularly in the selection of an attainment level for the criterion).

- What does "balanced natural populations" mean? What is the justification for quantitative criterion that represents that definition?

- Section 3.3 "Numeric Criteria Approaches" for estuaries is unbalanced in presentation and displays a distinct (and not defensible) bias toward EFDC/WASP/LSPC.

    - "Statistical relationships" is given relatively little coverage, with an initial focus on uncertainty analysis, (a strength of statistical models) while expressing concern over the additional uncertainty that can be introduced with covarying predictors. I mention this because there is essentially no coverage of uncertainty in the extensive mechanistic models section, ignoring the fact that there has never been a thorough uncertainty analysis for WASP and LSPC. Thus, statistical models are assessed (rightly so) in the document with concerns for uncertainty analysis, while this topic is virtually ignored for process-oriented models (more on that below).

    - Section 3.3.3 is so overwhelmingly more detailed than the sections on the other options (with much material from the user's manual that should simply be summarized here) that it implies that EPA has already chosen their approach. If so, this is premature, as major weaknesses are not addressed in the document.

    - Model evaluation (page 55) states "data from one or more different (from calibration years) years would be used to evaluate the performance of the model." For essentially all waterbodies, at best this is simply an examination of how the model changes under different hydrology (not different N,P concentrations in the inputs). Thus, if this is a test, it is a test of the model's response to hydrologic variation, not to changes in land use resulting in changes in nutrient input. After this evaluation, the model(s) would be applied to predict the "water quality responses that could be expected to occur in the absence of anthropogenic disturbance." This baseline conditions is far outside any model evaluation scenario, which means that the model will be applied to inform important decisions with no relevant testing of the model's ability to do that.

    - In Section 3.3.4, the document states "evaluate uncertainty associated with the model predictions." I know, and the EPA modelers know, that this cannot be done (if this means honest, through analysis of prediction uncertainty). At a minimum, the document needs to be re-written to clearly identify what uncertainties will be considered and what uncertainties will be ignored (and why).

- DPV
    - Why isn't cost a consideration for fraction allocation? For example, some upstream watersheds may have to deal with legacy nutrients in the soils.

See my comments above concerning mechanistic models; they apply here also.

Dr. David Schneider

1. **General Approach.**

a) General conceptual model. The model, as displayed in section 2.2, captures the key points. The model is then carried to selection of endpoints. Notably:

(1) Measures of the attribute or stressor
(2) Measures of exposure to the stressor
(3) Measures of ecosystems or receptor response.

This translation to measurement is consistent with the approach to environmental impact assessment in eastern Canada, which begins with Valued Ecosystem Components, identifies expected effects based on known causal links, and then identifies a triad of measurement, one from each of three categories:

(1) Monitoring continued presence of stressor (e.g. release of contaminant)
(2) Demonstrable effect of measured stressor level on biological rate (often lab assays)
(3) Monitoring ecosystem or receptor response (e.g. change in macrofaunal species composition).

The approach has been used with success in programs with high public visibility, including impact of offshore oil production, and effects of military activity on wildlife in Labrador.

My concern is that the tripartite approach is not carried through consistently in development of measures. A good example is the chapter on Florida coastal waters, were a single measure is used, receptor response, in category (3) only. Stressor levels (1) and exposure (2) are not considered. The result will be that (3) stressor response, will be unconnected to stressor levels and exposure.

In addition to (3) I would like to see a treatment of (1) stressor attributes (N, and P) and (2) stressor exposure. One possibility is an estimate of bulk release rates, based on known major sources (e.g. sewage, agriculture, etc). Are there reasons for omission of such estimates? There is a well developed literature on release of organic matter into the coastal zone, for example. It strikes me that effective institutional response to nutrient release in Florida waters will depend on knowledge of the major sources, and how these are changing over time.

b) Delineation of 4 Categories of Waters. The choices are consistent with my knowledge of estuarine, coastal, and offshore marine ecosystems. No science basis for three mile boundary appears, it is based on history?

In addition I have a few comments on Appendix B, Literature review.
Page B14. Phytoplankton assessment endpoints vague, not connected to prior table in text
Page B-33 Seagrass endpoint limited to a depth target for irradiance. Are others feasible relative to cost?
Page B-46 Coral dropped as intractable from point of view of (3) receptor response. No treatment of (2) exposure. Addressing (2) stressor exposure, strikes me as useful relative to a valued ecosystem component found nowhere else in continental US. The question is whether it is feasible.

3. Florida Coastal Waters

a) Data sources appropriate ? Sources in Section 2.4 are appropriate. Sections 4.1.1 and 4.2 do not mention appropriate date from 2.4 (meteorological data). This data, especially physical forcing at the sea surface, will be needed to interpret Chl, which is measured only at the sea surface.

a) Additional data sources ? Models estimates of mixing rates from wind stress at the seasurface would be useful. Estimates of bulk assimilation of N, P would be useful, if these exist.
Estimates of net release at the coast would be useful, based on river discharge, net outward flux through passes, and diffuse discharge in south Florida.

b) Chl endpoint appropriate? While it is appropriate, it addresses only (3) endpoint response. Chl is a state variable with multiple physical and biological forcers; and understanding of the underlying dynamics needs to be considered. Stressor levels and exposure to stressor will be needed to interpret the Chl endpoint. This is sufficiently important that a demonstration of non-feasibility will be required, for at least one measure of (1) stressor level and one measure of (2) exposure to stressor.

c) Appropriate application of remotely sensed Chl ?
Percentile approach (Section 4.2) better than use of means, which do not represent extreme events, which can be biologically important.

Note that the reference approach is weak unless spatial units can be assigned to N and P exposure categories. In the absence of such, the design is Before / After only, without being crossed with Contol / Impact.

In section 4.2.2, the calibration has weak statistical support (R^2 = 52%). It is not clear whether the spatial units for this regression are derived from spatial units of the same extent. Were averages within a pixel used? The skill of the model (slope = 0.85, s.e. not reported) needs to be quantified against new data. Use of remotely sensed Chl will depend on continuing calibration, give the statistical support. The comparisons in Figure 4.7 are based on means. Percentiles need to be considered, for consistency with percentile approach (Section 4.2).

3. Florida Coastal Waters (continued)
c) Appropriate application of remotely sensed Chl ?
Section 4.2.3. This appears to a binomial threshold, not a binomial test (a common statistical test, in the Neyman Pearson decision theoretic tradition).
Section 4.2.3.1. The implied model is log normal. Show that data are consistent with the model (e.g. normal after log transform). If not, find more appropriate model. A model is needed so that the percentile approach can be applied reliable. Note that log transform does not permit estimate of nutrient loading integrated over the coastal zone (multiplicative rather than additive).
Section 4.21.3.2. Define C_0.50 and C_0.75

A general comment concerning the 'binomial test.'
- this appears to be a way of quantifying exceedances at several levels, and thus is more informative that use of means. However,
- explanation is elusive. For example p127 reference to Section 3.3.2, where it is absent
- it is treated inconsistently across chapters.
- if quantifying uncertainty is important, then confidence limits are preferable to hypothesis testing, which appears to be the approach to uncertainty on p 117.

Perhaps some consideration should be given to quantification of risk that prevails in the medical literature (generalized linear model with logit link and binomial error).

Dr. Andrew Sharpley

General

The draft document "Methods and Approaches for Deriving Numeric Criteria for Nitrogen/Phosphorus Pollution in Florida's Estuaries, Coastal Waters, and Southern Inland Flowing Waters," describes three approaches in some detail, although there are still areas where further information and explanation is required. The strengths and weaknesses of each approach are given in Chapter 3. While it is commendable that this level of practical assessment was undertaken, there still remain questions as to the availability and level of detail of information required to adequately populate each approach.

On page 55, it states that a watershed model will be run with all anthropogenic sources removed to determine background TN and TP levels. More information and justification is needed to assure the users that the models being used can adequately accomplish this with a required degree of certainty. Most models have been developed on the premise of assessing and predicting fate and transport processes as a result of anthropogenic activities. Detailed validation of such approaches is needed, which mean calibration with non-impacted watershed loads. There are few non-impacted watersheds with pristine conditions that reflect baseline concentrations, in relation to determining water impairment. Reference sites in agricultural and urban settings are not common now and those that exist tend to be located in small watersheds. This begs the question of what level of nutrient loss / loading and biological response should be tolerated in consideration of U.S. food security, as well as cheap, clean water? While this is somewhat of a philosophical question, it does need to be addressed in derivation of numeric nutrient criteria by any method or approach.

Some consideration and assessment of the response variables used is recommended. For example, distinction is needed among nutrient forms that are of immediate availability to biological uptake (i.e., short-term bioavailability) and this growth response, such as nitrate-nitrogen, orthophosphate; compared with losses as particulate and organic forms of nitrogen and phosphorus (i.e., long-term availability). Some freshwater ecosystem studies have shown that chlorophyll *a* can be a function of grazing pressures rather than nutrient concentrations. For example, increasing nutrient concentrations in stream can increase the number of grazers, which can lead to a lower chlorophyll *a* concentration; i.e., a top-down regulation of primary production.

Model Approaches

It is not clear how the LSPC model approach will address certain objectives outlined in this document, such as measuring the effectiveness of site specific, highly variable Best management or conservation practices on both impervious and pervious surfaces, with an adequate degree of certainty (Section 3.3.3). The implementation of practices to decrease nutrient losses or inputs to surface waters (i.e., best or beneficial management practices - BMPs) can influence other factors that will affect biological response to nutrient loadings. For instance, riparian buffers are effective at removing sediment and sediment-bound nutrients (particularly phosphorus), as well as removing nitrogen by uptake and denitrification. However, they also provide shade and will influence stream water temperature and thereby the stressor-response relationship. This should be addressed in nutrient criteria development. In addition, the use of buffers for example, will influence the size of particulate or sediment in a stream or river that may affect the benthic population dynamics or species diversities. These direct and indirect effects and complexities should be captured in target setting or the evaluation of response to achieving target reductions.

More information is needed on the scale of resolution of predictive assessment by the LSPC model, which is stated to simulate water quality constituent wash-off using simple relationships to model nitrogen and phosphorus. It states that the model must "appropriately" represent the spatial and temporal variability of hydrological characteristics within a watershed and the rate at which nitrogen and phosphorus build up between rainfall events and wash-off, etc. More detail is needed to determine if this is an “appropriate” statement to make. Without this detail, there is a concern that this may be an over simplistic representation of reality, which will increase the uncertainty of model results.

Uncertainty

Throughout, uncertainty is briefly mentioned as being introduced because some environmental variables can covary with explanatory variable of interest. However, uncertainty issues related to numeric criteria should be described further and how they might influence the use and appropriateness of specific numeric criteria. For instance, can the coastal forces such as wind, currents, and tides, which influence coastal chlorophyll *a* dynamics together with nutrient loads, be accurately quantified (i.e., page 19)? It is essential that predictions explicitly state and detail the level of uncertainty inherent in those predictions and those predictions be “ground-truthed” (not “validated”) using site-specific data. The uncertainty among the various factors that are involved in the cause-effect relationship for a particular system of interest should be assessed.

The morphology of the aquatic system, habitat, spatial, and temporal relationships within the water body are all important in modifying the relationship between nutrient concentrations (both nitrogen and phosphorus) and observed endpoints. In fact these factors may dominate the cause-effect pathway so that nutrients are not important variables within the expected limits of the system. These factors need to be better documented, so that the uncertainty of the relationship can be reduced. A statistically significant stressor-response relationship can be derived that may represent only a small portion of the variability in the data. Relying solely on this relationship would result in a tremendous amount of uncertainty in the final criterion development.

Land Management

The Executive Summary states that “Numeric criteria will create environmental baselines that allow Florida to manage waters more effectively, measure progress, and support broader partnerships based on nutrient trading, best management practices (BMPs), land stewardship, wetlands protection, voluntary collaboration, and urban storm water runoff control strategies.” This may be an oversimplification of what actually might be able to be accomplished given the scale of model assessment, monitoring and response assessment compared with the scale at which management change occurs (i.e., field, farm, etc). In reality it will be very difficult to link source and impact, particularly at the scale where management in conducted on the landscape. Comparisons are made and similarities drawn with the pattern of nitrogen transport in the Mississippi / Atchafalaya River Basin as some justification for the current estimates in southeast U.S. streams (page 131). However, it is not clear if this is a function of watershed properties (i.e., reality) or model formulation. It is also unclear how the two widely differing spatial scale processed-based transport, can be compared or equated from information in this report.

The importance of legacy nitrogen and phosphorus effects from past management and natural sources should not be understated or overlooked. The legacy effects of past management appear to have been minimized by several statements throughout the document (e.g., top of page 40 in reference to Huang and Hong, 1999). More recent work by Reddy on nutrient storage in Florida ecosystems for NRC suggests otherwise.

How will hereditary or legacy losses or inputs of nitrogen and phosphorus to water bodies be considered and accounted for in such an empirical approach? This begs the next question facing water resource managers who set targets for nutrient loss reduction, that if no water quality improvement or indicator biological response is seen, are the targets / criteria too low, are legacy nutrient inputs an increasingly significant contributor, or how long does it take dynamic ecosystems and watersheds to respond to change? How will continued legacy or hereditary inputs of stressor inputs (nitrogen and phosphorus) be

distinguished from management change-related decreases? Internal recycling of nutrients can mask water quality improvements brought about by nutrient loss reductions brought about by land management changes.
It is implied that continued monitoring and assessment of concentrations and biological response in all designated ecosystems is taking place. It is not clear if this is being done, at what scale and how. This relates to the issue of measuring biological indicator change to reducing nutrient inputs to water bodies. For example, clearly there will be a water quality / water body response to change in nutrient inputs wither through point or nonpoint source efforts. How will hereditary or legacy losses or inputs to waters of nitrogen and phosphorus be considered and accounted for the described approaches? This begs the next question of water resource managers who set targets for nutrient loss reduction to meet numeric criteria, that if no water quality improvement or indicator biological response is seen, are the criteria (or targets) too low, are legacy nutrient inputs an increasingly significant contributor, or how long does it take dynamic ecosystems and watershed to respond to change?

Dr. Alan Steinman

a) *EPA has introduced a general conceptual model in Chapter 2, including the selection of assessment endpoint and indicator variables. What is your perspective of the general conceptual model?*

The overall structure is generally sensible, but I have 4 specific comments on the model:

1. I wonder if it makes sense to first have a classification of dystrophic or colored waters vs non-colored waters? The use of chl *a* as a response variable is questionable in the former class of waters, and as a consequence, the utility of the model is limited. In this case, there would be 2 conceptual models developed, one that does not use chl *a* as the key response variable (for naturally colored waters) and the current model that does use chl *a*.
2. It is unclear to me why chl a concentration is a target for the PPL community structure? This seems like a red herring—the real target here is a desirable algal community based on taxonomic composition, irrespective of biomass. The biomass target is dealt with in the other targets.
3. The focus on chl a, while understandable, oversimplifies potential responses—the benthic autotrophs with root systems (e.g., seagrasses) meet much of their nutrient demand from the sediment, not the water column. This process is not accounted for in this model. Might it make sense to have benthic and pelagic models, or does this make the conceptualization too complicated?
4. This may be outside the scope of our Charge, but I find the use of the word "Objective" as the final outcome objectionable (couldn't help myself there). Objectives trigger these conceptual constructs, not what concludes them. Wouldn't societal values be a better term?

b) *EPA has delineated the State of Florida into 4 general categories of waters—Florida estuaries, Florida coastal waters, South Florida inland flowing waters, and South Florida marine waters—for purposes of considering approaches to numeric nutrient criteria development . Are these categories appropriate and scientifically defensible? (Note that the details of segmentation of waters within these categories is addressed in subsequent charge questions.)*

These categories make sense given the overall approach. I have only 2 comments, which are elaborated on below: 1) consideration might be given to separating out the Caloosahatchee and St. Lucie estuaries from the other Florida estuaries, given their unique (i.e., strong human influence) hydrological relationship to Lake Okeechobee—while nutrients clearly influence their biota, salinity (or lack thereof) also plays a stronger role than what is typically found in other Florida estuarine systems (cf. Kraemer et al. 1999 Estuaries 22: 138-148; Doering et al. 1999 Florida Scientist 62: 89-105; Steinman et al. 2002 Conservation Ecology 6(2):17); and 2) the category of South Florida inland flowing waters seems to be a grab bag for waters that don't fit anywhere else—the land of misfit hydrology, if you will. It would be preferential to have a strong scientific rationale for this classification, as opposed to a default category.

2

2. Estuaries

a) Are the data sources identified appropriate for use in deriving numeric criteria in Florida's estuaries (as discussed in Sections 2.4 and 3.2)? Is the SAB aware of additional available, reliable data that EPA should consider in delineating estuaries or deriving criteria for estuarine waters? Please identify the additional data sources.

The data sources appear appropriate, at least to initiate the process. I have identified other data sources throughout my responses to the charge questions, in the appropriate locations.

b) Are the assessment endpoints identified in Sections 2.3 and 3.2 (healthy seagrass communities; balanced phytoplankton biomass and production; and balanced faunal communities) appropriate to translate Florida's narrative nutrient criterion (as cited above) into numeric criteria for Florida's estuaries, given currently available data? Does the SAB suggest modification or addition to these assessment endpoints? A literature review of endpoints considered can be found in Appendix B.

Again, I provide alternative or complementary approaches throughout my responses. I believe EPA's approach is a good starting point. However, I have the following concerns:

☐ using chlorophyll *a* as a proxy for phytoplankton species composition—it is too crude an indicator to provide useful information for some situations

☐ while phytoplankton production is listed as an endpoint, in fact biomass (via chl a) is the endpoint, not the metabolic process; chl a gives an index of net production, so the use of the term should be qualified

☐ partitioning the estuaries that are hydrologically linked to Lake Okeechobee from the other estuarine systems, because salinity influences seagrass health as much, if not more so, than nutrients in the St. Lucie and Caloosahatcee

☐ the influence of colored dissolved organic matter, in lieu of chlorophyll, as a light extinction factor

c) EPA describes potential approaches in Section 3.3 (reference conditions, stressor response relationships, and water quality simulation models) for deriving numeric criteria in Florida's estuaries. Compare and contrast the ability of each approach to ensure the attainment and maintenance of natural populations of aquatic flora and fauna for different types of estuaries, given currently available data?

The analysis did by EPA was quite thorough. My general thoughts on each approach:

I believe the reference condition approach has the greatest heuristic value because of its scientific transparency and defensibility, and it is also understandable to the general public; unfortunately, it has limited applicability if the data are not available. Of course, even with good reference condition data, confounding factors may preclude this approach's ability to protect the biota. For example, there may be usable reference condition data for seagrass distribution in the Caloosahatchee Estuary, but changes in the regulation schedule of Lake Okeechobee may prevent their restoration to reference state because of an altered salinity regime.

3

The stressor response relationship also benefits from being readily understandable and allows uncertainty to be both statistically derived and perhaps more importantly, graphically represented. However, one must be wary in its use, as it can lead to false expectations, especially with linear (or close to linear) relationships. It is critical to emphasize that these relations must not be extrapolated beyond the existing data, as there is no way to know if the response variable is near a tipping point, beyond which its behavior may no longer demonstrate a linear relation with the stressor. This is particularly acute with nutrients, where bloom formations rarely demonstrate a linear relation with N or P. However, by including confidence intervals in these relations, it is possible to graphically show what levels the stressor(s) needs to be held to in order to protect the biota.

The coupled simulation models provide much more flexibility in studying system behavior, and therefore have intuitive appeal, especially in predicting future change. However, they run several risks: 1) no matter how useful, they ultimately are simplifications of the ecosystem, and therefore are inherently wrong—most ecologists, other than pure empiricists, are not particularly bothered by this limitation, but it can present potential problems when their use is explained to the general public; and 2) model calibration may improve model performance, but as just noted, that does not mean the model is accurately representing system behavior—it would be much more valuable to include a multi-faceted approach (if resources allow), where models are used in conjunction with observational and experimental approaches to identify the appropriate nutrient criteria.

3. Coastal Waters

a) Are the data sources identified in Sections 2.4, 4.1.1 and 4.2 appropriate for use in deriving numeric criteria in Florida's coastal waters? Is the SAB aware of additional available, reliable data that EPA should consider in delineating coastal waters or deriving criteria for coastal waters? Please identify the additional data sources.

This is outside of my area of expertise, and I cannot answer knowledgably.

b) Is the assessment endpoint identified in Section 4.2 (chlorophyll-a to measure balanced phytoplankton biomass and production) appropriate to translate Florida's narrative nutrient criteria (described above) into numeric criteria for Florida's coastal waters, given currently available data? Does the SAB suggest modification or addition to this assessment endpoint?

Chlorophyll *a* has some clear advantages as an endpoint: data availability and comparability (but see below), ease of measurement, and understanding of its ecological role. However, as noted above, it is a crude measurement that provides no information on species composition and its level is influenced by ambient light and nutrient levels, so one must be careful when comparing concentrations across time or space. In addition, it is a poor surrogate for production, as it really represents **net** production (something not addressed in this document) and as a state variable, does not represent a metabolic process.

More rigorous endpoints for balanced phytoplankton biomass might focus on a percentage of cyanobacteria not to exceed some level, or even a not-to-exceed percentage of certain

cyanobacteria or dinoflagellates species. Similarly, a more rigorous endpoint for phytoplankton production might be an actual rate of primary productivity (either in terms of carbon fixed or DO evolved). Obviously, the endpoints identified in this paragraph lack the data availability and ease of data collection that characterize chlorophyll a, so ultimately, they may not be practical choices. However, it is important to recognize the limitations of chlorophyll *a*.

c) Does the approach EPA describes in Section 4.2 appropriately apply remote sensing data to ensure attainment and maintenance of balanced natural populations of aquatic flora and fauna in Florida's coastal waters? If not, please provide an alternate methodology utilizing available reliable data and tools, and describe the corresponding advantages and disadvantages.

I don't have expertise in the field of remote sensing, so at least superficially, this approach seems practical. Clearly, field-based alternatives to assess coastal chlorophyll levels would be expensive and time-consuming. I do have two comments though based on my reading of the section:

1) p. 83, para 1: I assume that the "flagging" of the 1-wk prior to bloom data points is akin to removing the points from the analysis? This is not clear, and should be explicit. I recommend that they are, in fact, removed as those antecedent conditions leading up the

bloom are likely not representative of desired “reference conditions”.
2) p. 84, para 2: There appears to be a bit of data cherry-picking at work in this paragraph, as the text ignores the lack of correspondence between the satellite and field data during 2001-2004 (significant decline in the field data, but relatively flat [on average] levels for satellite data). Shouldn’t this assessment focus on areas of disagreement, as well as areas of agreement?

4. South Florida Inland Flowing Waters

a) Are the data sources identified in Section 2.4 and 5.4 appropriate for use in deriving numeric criteria in South Florida’s inland flowing waters (as discussed in Chapters 2 and 5)? Is the SAB aware of additional available, reliable data that EPA should consider in delineating or deriving criteria for South Florida’s inland flowing waters? Please identify the additional data sources.

These data sources are certainly the most logical to start with. EPA may want to look into datasets potentially available from the local water/drainage districts (not water management districts), such as Lake Worth and Loxahatchee, as well as from agricultural interests that border the canals (e.g., US Sugar) although those data tend to be proprietary.

b) Are the assessment endpoints identified in Section 5.4 (balanced faunal communities, i.e., aquatic macroinvertebrates, and balanced phytoplankton biomass and production) appropriate to translate Florida’s narrative nutrient criteria (described above) into numeric criteria for South Florida’s inland flowing waters, given currently available data? Does the SAB suggest modification or addition to these assessment endpoints?

5

I think macroinverts and water column chl *a* are reasonable endpoints (but see my other comments regarding chl); as noted in my general comments, I also recommend EPA consider the use of algal community structure (both planktonic *and* benthic) as endpoints. Periphyton community structure has been shown to be a sensitive indicator of phosphorus impairment in South Florida (cf. McCormick et al. 1996 JNABS 15: 433-449; McCormick and O’Dell 1996 JNABS 15: 450-468; McCormick et al. 2002 pp 83-126 in: The Everglades, Florida Bay, and Coral Reefs of the Florida Keys), and although it is more labor-intensive and requires more technical expertise than measuring chlorophyll, it is a more definitive and rigorous endpoint for many systems in South Florida. To my knowledge, it has not been used as an indicator in these canal systems, but I believe it is worth exploring if any work has been done in this area.

c) EPA describes two approaches in Section 5.4 (reference conditions and stressor-response relationships) for deriving numeric criteria in South Florida inland flowing waters. Compare and contrast the ability of each approach to ensure attainment and maintenance of balanced natural populations of aquatic flora and fauna in different types of flowing water or geographical areas, given currently available data?

Both approaches have merit, and absent more information on their respective performance, it would be premature to assume one performs better than another.

As for the reference condition approach, the LDI approach has intuitive appeal; similar metrics have been applied elsewhere and seem to work well, at least at a generic level. I assume the 100 m buffer is based on Lane’s original assessment for the flat landscape of Florida; this should be noted explicitly in the report, as I suspect some readers may be curious. As I recall, Lane’s Ph.D. was not focused on South Florida specifically, so I am curious if any additional work has been done to validation the application of LDI (especially the coefficients) to this region? If not, I recommend caution in its use, at least until the validation can be performed, as it is unclear if the endpoints for these inland flowing waters behave similarly as for depressional wetlands (*full* *disclosure: I served as the Associate Editor for 2 submissions by Lane et al. back in 2003-2004 to JNABS on this LDI that ultimately were not published—I believe that submission did serve as the basis for the EMA paper*).

The annual average concentration exceedance approach also has intuitive appeal, but the current description does not address magnitude or duration at all. If this approach is to be adopted, the probability functions in Table 5-2 will need to be amended to account for these other components.

The stressor-response relationship has a solid conceptual foundation and has been used successfully in other systems. However, based on the available data as shown in Figs. 5-11, 5-13, and 5-15, there are no obvious break points at which impairment can be determined, at least for planktonic chl *a* concentration. Perhaps different ecosystem responses may respond in a more "cooperative" manner, but if not, this approach may have limited utility.

Finally, just a general observation that people unfamiliar with south Florida waterways may have the wrong impression of the term "inland flowing waters". EPA may want to explain or define this term, preferably near the start, using comparative current velocity data to show that flowing

waters can be very sluggish compared to what many think of when the term "flowing waters" is used.

5. South Florida Marine Waters

a) Are the data sources identified in Section 2.4 and 5.5 appropriate for use in deriving numeric criteria in South Florida's marine waters (as discussed in Chapters 2 and 5)? Is the SAB aware of additional available, reliable data that EPA should consider in delineating or deriving criteria for South Florida's marine waters? Please identify the additional data sources.

The SERC monitoring network seems extensive with 350 sites, although it would have been informative to include a map of those sites to depict their distribution. Do they extend from the Gulf around to the Atlantic, or are they concentrated in Florida Bay?

b) EPA describes two methods in Section 5.6 for using a reference condition approach for deriving numeric criteria in South Florida marine waters (least-disturbed sites or bionomial test). Compare and contrast the ability of each approach to ensure attainment and maintenance of balanced natural populations of aquatic flora and fauna in South Florida marine waters, given currently available data?

The approach for deriving numeric criteria appears reasonable, but my above concerns for exceedances still remain; regardless of whether a single-number or two-number threshold is applied, neither address the issues of duration and magnitude from what I can infer.

I appreciate the example provided in section 5.6.1.3 to illustrate the approach. You may want to consider presenting the TP data in ppb instead of ppm so you don't need to use so many significant digits (which can be distracting).

6. Downstream Protection Values for Florida Estuaries and South Florida Marine Waters

a. Are the methods EPA is considering for deriving downstream protection values (DPVs) for estuaries (excluding marine waters in South Florida) as described in Section 6.1-6.4 appropriate to ensure attainment and maintenance of downstream water quality standards, given available data? Please describe additional approaches and their advantages and disadvantages that EPA should consider when developing numeric criteria to protect these downstream estuarine waters (excluding marine waters in South Florida), given available data?

Conceptually, the overall approach has intuitive appeal, but as always, the devil is in the details. My main concern, at least based on the available information, is in the calculation of the first order decay rates. Given the much slower flow of most of the South Florida inland waters, the decay coefficients may be much higher than what has been calculated by Hoos and McMahon for TN and Garcia et al. for TP. I have not seen their datasets, so their calculations may in fact include data from these streams; if so, my concern is less of an issue. If not, I recommend that some preliminary research be conducted on selected inland flowing waters in South Florida to validate these decay rates.

7

b. Are the methods that EPA is considering for deriving downstream protection values (DPVs) for marine waters in South Florida as described in Section 6.5 appropriate to ensure attainment and maintenance of downstream water quality standards, given available data? Please describe additional approaches and their advantages and disadvantages that EPA should consider when developing numeric criteria to protect downstream marine waters in South Florida, given available data?

I am not sure I fully understand this approach, so take my comments with a grain or two of salt:

1) I believe there is value in applying both the constant and first-order decay coefficients in order to get estimates of uncertainty in the criteria. This may make it difficult, if not impossible, to assign a single nutrient standard, but any single number is fraught with issues anyway. Providing a range of possible protective values has some appeal.

2) Computationally, I understand how the nutrient concentrations decline (with distance from their source) as they reach the receiving water body (e.g., Fig. 6-7). However, in many watersheds, the areas near the receiving water body are heavily developed, and therefore nutrient concentrations end up increasing, not decreasing, because of stormwater runoff and limited ability for ecosystem assimilation (e.g., Groffman et al. 2003 Front Ecol. Environ. 6: 315-321; Steinman et al. 2006 Arch. Hydrobiol. 166: 117-144). I realize these model results are strictly for illustrative purposes, but I hope these kinds of land use issues are taken into account.

3) Finally, and this may be well outside the purview of the Panel's scope, I wonder if EPA will consider including mitigation practices in their nutrient criteria setting approach. For example, could a modeling scenario be included that allows users to implement BMPs in the watershed, and thereby reduce nutrient loads?

General thoughts:

1) How is "imbalance in natural populations" defined? It is such a vague term, which was undoubtedly intentional, but it makes it difficult to assess imbalance without a standard to compare against.

2) There is no explicit mention of future potential impacts, such as climate change or predicted land use change. Should that be considered or do we restrict ourselves to present threats only?

3) p. 49: Not clear to me how chl a can substitute for species composition changes?

4) Just an fyi that we conducted a review of WAM last year for the SFWMD, as part of a phosphorus-related project on the Kissimmee River, wherein we identified its limitations.

5) p. 69: It is not clear to me if the IRL in Table 3-6 is inclusive of the Loxahatchee Estuary—if not, then the Loxahatchee should be listed separately, with good seagrass and water quality information available from the Loxahatchee River District (see studies by Dent and co-workers) and the SFWMD (cf. Noel et al. 1995 JAWRA 31: 21-32).

8

6) p. 74: The assignment of reach group is logical, but it is unclear if these assignments are *a priori* or post hoc? This is not just a philosophical question—it has bearing on scientifically defending the modeling approach. For example, an *a priori* approach could

be based on best professional judgment, whereas the ad hoc approach may depend on statistical groupings of stations, as determined by cluster analysis. The decision framework regarding how the reach groups are determined needs to be explicitly stated.
7) p. 75: Unclear if "detachment" refers to desorption of nutrients from sediment particles, or some other phenomenon.
8) p. 80, Table 4-1: I assume column 4 is the "corrected" chl a data? This is implied, but perhaps should be explicitly stated.
9) p. 97-98: It is worth noting that the 10 ppb criterion was largely based on responses by periphyton to phosphorus. Interestingly, there is no mention of this biotic community as a response variable in this document, which is somewhat surprising given the amount of research and analysis this relationship has received.
10) p. 99: Need to be more explicit about how the Cl threshold of 1500 ppm will be calculated—is it a daily, monthly, annual geometric mean? Sampled at baseflow or during storm events, or both? Will sampling control for tidal activity? Will samples be depthintegrated or just near-surface collections? Lots of questions and little detail.
11) p. 123, section 6.2, last sentence: Presumably, this one protective loading rate would be reflect the most conservative rate for the entire estuary? This should be explicitly noted.
12) The report is generally well-written but the document should be checked for typos.

Dr. Andrew Solow

I will largely confine my preliminary comments to statistical issues arising in the report. I do, however, have a couple of more general comments. First, the document refers repeatedly to 'balanced communities' as an assessment endpoint. There is a need for more discussion about what this means. Second, in places, the document appears to suggest that conditions in waters affected by human activities can or ought to be returned to pristine conditions – even conditions pre-dating human settlement. It is nearly a tautology that this cannot reasonably be expected to happen. This is not to say that human activities are incompatible with reasonable assessment endpoints.

On the statistical side, the document refers to two broad approaches. The reference condition approach involves developing criteria by examining conditions in some collection of sites. In some cases, the reference sites are those that are identified as 'unimpaired'. In other cases, the reference sites represent a sample of all sites. The underlying logic here is different for these two possibilities. The logic underlying the former is that, very roughly speaking, the criteria will make the impaired site look like an unimpaired site. The logic underlying the second is that, very roughly speaking, no impaired site should resemble – or move toward – the worst impaired sites (defined by an upper quantile of a measure of impairity). It is worth thinking these options out. The technical details of these approaches – based on the quantiles of the log measurements or the 'binomial' approach – need to be worked out in more detail.

The second statistical approach described in the document involves developing regression models relating either quantitative endpoints or indicators to 'stressors'. I was struck by the relative simplicity of the examples in the document. Particularly in light of the cost and complexity of approaches based on coupled numerical hydrological-water quality models, the statistical approaches are well below the usual standard in environmental statistics. There is room for improvement.

Finally, there is a connection between the reference condition approach and the approach based on regression modeling of the impact of stressors on indicators. Under the reference condition approach, the appropriate level of stressor is determined by examining its level at sites with favorable assessment endpoints (presumably corresponding to favorable levels of indicators). In contrast, the regression approach uses the full range of sites – impaired and unimpaired – to determine stressor levels that provide favorable values of indicators. To put it another way, the reference condition approach focuses only on the favorable end of the spectrum of outcomes while the regression approach considers the full spectrum. What the reference condition approach is missing needs to be thought through carefully.