Comments on Nutrient Criteria for Florida's Estuarine and Coastal Waters

Frank Marshall to: Stephanie Sanzone 01/30/2011 01:49 PM

Ms. Sanzone:

Attached please find comments on the EPA document 'Nutrient Criteria for Florida's Estuarine and Coastal Waters, and Southern Canals - Methods and Approaches for Deriving Numeric Approaches'. My comments are in the form of an implementation of the suggested approach for marine waters as provided in Section 5.6 of the EPA document. This implementation identified several issues with the described approaches that can only be seen through an implementation exercise as attached. Please include this example in the comments that will be reviewed by the Science Advisory Board for this effort.

Please also note that I am a member of the Florida Department of Environmental Protection (FDEP) Marine Technical Advisory Committee if a disclosure is needed.

Thanks.

Frank Marshall PhD
Cetacean Logic Foundation, Inc.
New Smyrna Beach, Florida

An Application of the EPA Approach to Setting Numeric Criteria for Nitrogen/Phosphorus Pollution in Florida's Estuaries and Coastal Waters

Prepared by Frank Marshall, FDEP Marine TAC Member

January 28, 2011

Introduction

In the November 17, 2010 *Methods and Approaches for Deriving Numeric Criteria for Nitrogen/Phosphorus Pollution in Florida's Estuaries, Coastal Waters, and Southern Inland Flowing* Waters the US Environmental Protection Agency (EPA) has proposed two approaches for deriving numeric nutrient criteria for marine waters of South Florida (EPA, 2010). Both approaches utilize a type of reference condition. One approach uses average conditions from reference water bodies that are considered to be 'least disturbed'. The other approach uses descriptive statistics for a distribution of existing nutrient data, and the statistics are used in a binomial distribution evaluation. Both average/median conditions are evaluated as well as higher nutrient values that may not occur as frequently but are potentially damaging to the ecosystem. These criteria are considered to inherently protect a coastal water body because they are associated with a site where the biota has been found to be, in general, healthy.

The least-disturbed approach uses annual geometric means as is described in EPA (2010) for both marine and inland flowing waters. The proposed criteria are comprised of three components, magnitude, frequency, and duration. The magnitude component is a test that utilizes a distribution of annual geometric means from a reference set of data, with frequency and duration of exceedance (how often the criteria are exceeded over a set period of time) set at the 75th percentile (P75) value of the annual geometric mean distribution. Confidence in the evaluation (not the data) is incorporated through probability of exceedance of P75 with 80-90% confidence considered an acceptable level of risk for this evaluation. Using a five year period of evaluation, this means that frequency of exceedance of the P75 value from the distribution of annual geometric means cannot exceed two exceedances in three years or three exceedances in five years. A concern is expressed by EPA about the adverse effect of two back-to-back exceedance years.

EPA also presented an approach that did not differentiate between least disturbed and disturbed water bodies. This approach pools all of the data from a water body or sub-area of a water body, and assumes that the data distribution includes ambient conditions across a disturbance gradient. Because disturbed and least disturbed sites are included, a lower percentile would be specified as the criteria needed to sustain a healthy ecosystem. In the approach document EPA does not suggest an acceptable percentile.

This document presents an implementation of the numeric nutrient criteria information in EPA (2010) by evaluating data from a south Florida estuary. This is one interpretation of the EPA proposed approach and may or may not be the manner in which the information in EPA (2010) was intended to be used. If it is not as intended, then this may be an indication that additional clarification as to use is needed by EPA.

Data

The data used for this comparison were taken from a dataset for a south Florida estuary compiled by the author for use with a water quality box model that was developed to evaluate changes to nutrient inputs due to urbanization of existing agricultural land. The data have also been previously used recently to evaluate one of several suggested approaches being examined by the Florida Department of Environmental Protection (FDEP, 2010) in a parallel evaluation of "straw man" numeric nutrient criteria developed by FDEP.

The data are a combination of Total Phosphorous (TP) data from several sources that were combined into a time series of monthly average values from October 1993 through March 2007. These data were collected from 30+ sites and analyzed using the same protocol. The land use surrounding some of the sites would be considered least disturbed and land use around other sites would be considered disturbed. For simplicity, these data were pooled spatially into average monthly measurements, representing a water body where spatial coverage of monitoring sites is limited as is the case in some Florida estuaries.

Summary statistics for the full dataset are presented by Table 1. The statistical parameters used for the analysis were computed mostly by SAS ©, but can be computed by Excel © data analysis applications. The monthly time series for TP is presented in Figure 1. For the final water year of the time series (2007) data are available for only the first 7 months of that year, representing a case where there are missing values for a particular year of evaluation. Figures 2 and 3 present frequency of measurements in bins for non-transformed and natural log (LN) transformed data, respectively. The non-transformed data are skewed as is typical for nutrient data. The shape of the frequency distribution for the LN data approximates a normal distribution.

Table 1. Summary statistics for the monthly TP data used for this evaluation, October 1993 through March 2007.

Mean	6.15
Geometric Mean	5.42
Standard Error	0.33
Median	4.89
Standard Deviation	4.24
Variance	17.96
Minimum	2.43
Maximum	39.96
Count	161

Figure 1. Spatially-averaged TP data from a Florida estuary from monthly grab samples.

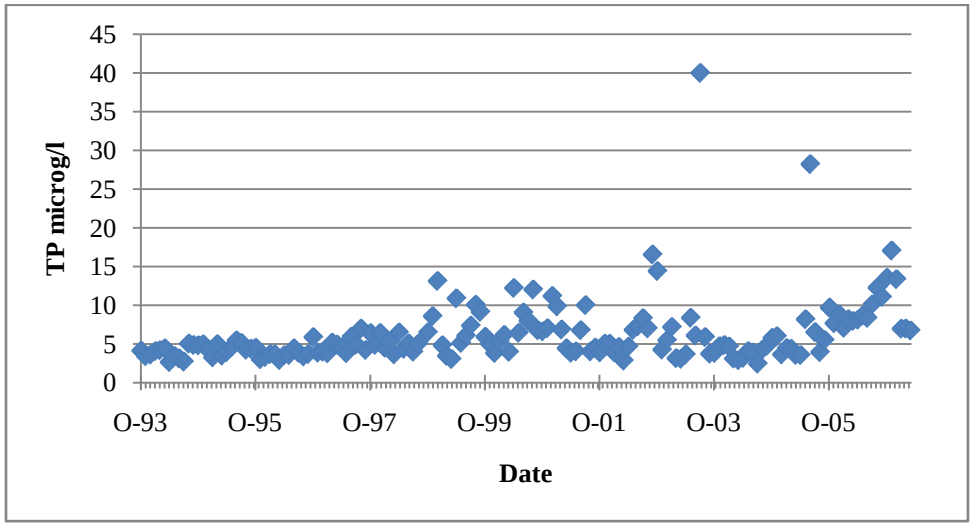


Figure 2. Frequency chart for non-transformed TP data. TP values are maximum value for a particular bin.

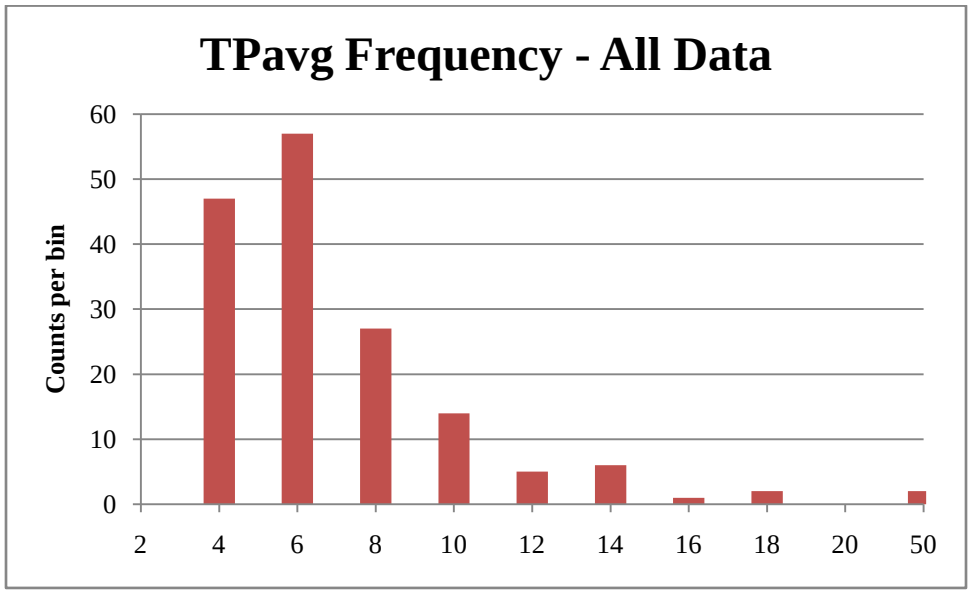
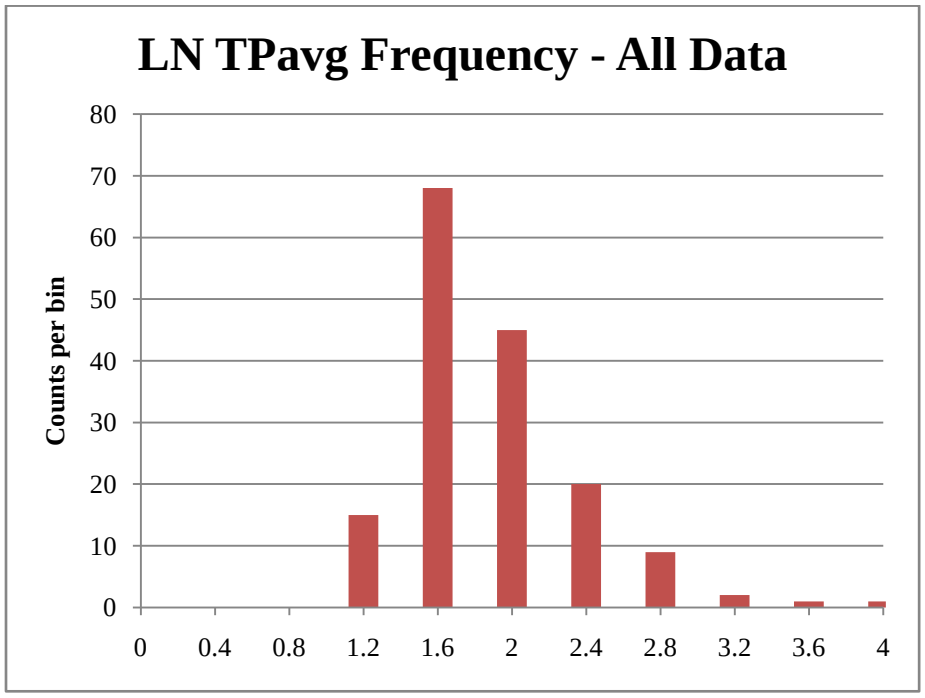


Figure 3. Frequency chart for natural log-transformed TP data. TP values are maximum value for a particular bin.



Analysis

The TP data were divided into two data sets using an October-to-September water year, as follows:

1. A Base Period from October 1993 through September 2002 (9 years), and
2. An Evaluation Period from October 2002 through September 2007 (5 years).

Tables 2 and 3 present summary statistics for the Base Period and the Evaluation Period TP data, respectively. The range of TP values in the overall dataset is wide (Table 1), and Figure 1 shows that, in general, the low TP values are 3 -7µg/l and the high TP values are about 10 - 15µg/l, though there are 2 values above 25µg/l. The mean and median values for the overall data (Table 1) indicate that, in general, this estuary is an oligotrophic system.

Two evaluations were made: an average annual concentration (long term) evaluation, and an upper limit evaluation using a binomial test intended to maintain existing conditions. The average annual concentration evaluation as specified by EPA in Section 5.6.1 of EPA (2010) is a test that utilizes a distribution of annual geometric means from, in this example, a nine-year reference set of data. The magnitude criteria is the P75 value with frequency and duration of exceedance of the P75 value over the five-year evaluation period set at two exceedances in three years or three exceedances in five years establishing an 80-90% level of confidence.

Table 2. Summary statistics for the Base Period TP data.

Mean	5.45
Geometric Mean	5.05
Standard Error	0.24
Median	4.81
Standard Deviation	2.44
Minimum	2.64
Maximum	16.53
Count	107

Table 3. Summary statistics for the Evaluation Period TP data.

Mean	7.58
Geometric Mean	6.23
Standard Error	0.85
Median	5.93
Standard Deviation	6.27
Minimum	2.43
Maximum	39.96
Count	54

For the evaluation period the geometric mean of all five years of data (N=43, some missing values) is 6.21 µg/l. The annual geometric means for these five years of data range from 3.79 to 10.0 µg/l. The annual geometric mean for 2006 (8.88 µg/l) and 2007 (10.0 µg/l) are greater than the P75 value for the annual geometric mean distribution for the base period data (6.15 µg/l). On this basis the two-exceedance-in- three year criterion is exceeded, but the three-exceedance-in-five year criterion is not. However, the lower 90% one-sided confidence limits for both years (2006 = 6.78 µg/l; 2007 = 7.05 µg/l) are less than the base data upper 90% one-sided confidence limit. This means that, when the variability in the two data sets is included in the analysis, it cannot be concluded with a 95% level of confidence that an exceedance occurred in either 2006 or 2007.

For a second evaluation, both sets of data were divided into 3-year rolling window periods, and the P50 and P75 were computed for each 3-year rolling window period. The base period data were used to develop criteria which were compared to statistics from the full evaluation period to determine if the TP criteria from the base period were exceeded during any of the rolling average periods for the evaluation period.

Table 4 summarizes the values computed for the 3-year rolling window analysis. The period 2000-2002 has the highest P50 value and the second highest P75 value and the statistics for this period were selected as the base period criteria for comparison to the evaluation period data.

When the P50 values for each 3-year period in the evaluation period data are compared to the same values for the 2000-2002 base data there is one 3-year exceedance period, 2005-2007. When the P75 values for each 3-year period in the evaluation period data are compared to the same values for the 2000-2002 base data there are two 3-year exceedance periods that exceedance periods, 2004-2006 and 2005-2007. However, when 90% one-sided confidence intervals are computed using coverage similar to that use for confidence intervals around the mean, there are no exceedances (Table 5).

Table 4. A summary of the data used to analyze the TP data for exceedances according to the EPA approach (binomial test to maintain current conditions). Bold text indicated the criteria used.

Period	Year	P50	P75	One-side 95% CI	
Base	94-96	4.00	4.56		
	95-97	4.24	4.95		
	96-98	4.50	5.75		
	97-99	5.05	6.35		
	98-00	6.15	8.25		
	99-01	6.10	8.00		
	00-02	***6.50***	***8.10***	***6.50-7.35***	***8.10-9.0***
Evaluation	01-03	4.95	7.05	lower	lower
	02-04	4.64	6.75	lower	lower
	03-05	4.55	6.10	lower	lower
	04-06	6.25	8.70	4.86-6.25	7.31-8.7
	05-07	8.10	9.95	6.47-8.1	8.3-9.95

Table 5. Summary of the 3-year rolling window exceedance analysis for the evaluation period.

Period	P50 Exceedance?	P75 Exceedance?	P50 CI Exceedance?	P75 CI Exceedance?
01-03	no	no	no	no
02-04	no	no	no	no
03-05	no	no	no	no
04-06	no	yes	no	no
05-07	yes	yes	no	no

Discussion

There are a number of ways to express numeric nutrient criteria and to compare the criteria with monitoring data for the purpose of determining the nutrient status of an estuary or coastal water. The available data show that the estuaries in Florida are quite variable and one set of values for numeric nutrient criteria is not going to be applicable to all Florida marine waters. Recognizing this variability, EPA has proposed setting separate numeric nutrient for South Florida.

A cursory comparison of the summary statistics for the Base and Evaluation Periods (Tables 2 and 3) indicate that these datasets may not be similar. In fact, when the annual average concentration and binomial test evaluations are made using only the applicable percentile values without confidence intervals, there are criteria exceedances for 2006 and 2007, as well as the 2004-2006 and 2005-2007 three-year periods. However, when the variability of the data used to compute these statistics is included, there are no exceedances.

The reason for these seemingly contrasting results is in the small number of values used for the evaluations. For the average annual concentration evaluation the number of values used to compute the P75 value for the base period is 9. The number of values used to compute a geometric mean for the data of any year in the evaluation period to be compared with the P75 value of the annual geometric means is 12. For the three-year rolling window evaluation the number of values used to compute the P50 and P75 values is 36. All of these N values are small when the variability of data is taken into account. Therefore, when one-sided confidence intervals are computed the relatively large variability of the small number of values causes the range of confidence interval values to be large, and it is difficult to determine if the criteria have been exceeded.

For the annual concentration evaluation, adding one annual geometric mean value to the base data requires collecting a year of data, regardless of the number of sites being monitored. Many years of data collection will be required before the number of annual geometric mean values is capable of reducing the range of the confidence intervals for the P75 value. For a single year annual geometric mean, additional monitoring sites will provide a better estimate of the geometric mean. For the 3-year rolling window, additional monitoring sites will provide a better estimate of the geometric mean and perhaps a narrower range for the confidence intervals. Alternately, for all evaluations the level of risk could be reduced from the 95% level of confidence. However, each of these improvements to the datasets except relaxation of the level of confidence requires additional monitoring. The annual concentration evaluation appears to be particularly hampered by this constraint.

It is important to note that there are a number of Florida estuaries and coastal waters with limited monitoring data as was used for the example above. If the data variability based on a limited sample size is large and small sample size is suspected as a problem, a period of monthly sampling may need to be initiated until the true annual variability in the data can be better interpreted. If a water body has no long-term or short-term exceedances during a particular year, the data from these periods should be incorporated into the base period to strengthen the base period data, so long as the supporting ecological data do not indicate a worsening of eutrophic conditions even though the water quality data do not indicate the same.

In addition, there have been other approaches proposed for setting the threshold for frequency and duration criteria (Briceno et al., 2010) which are alternatives to the proposed EPA approaches. There is merit to an approach to setting nutrient threshold criteria that utilizes a

reasonable interpretation of the productivity response to increasing levels of nutrients in the water rather than the nutrient data itself. Nutrient bio-utilization may be subject to factors such as N:P ratio, etc. Consideration should be given to adoption of the Briceno et al., 2010) approach.

Review of the examples in Briceno et al. (2010) indicates that the threshold values developed from the Z-cusum approach may be lower than the adjusted 75th percentile value and would provide additional protection against nutrient-induced eutrophication. Support for using this or a similar approach can be found in the Precautionary Principle: "Where an activity raises threats of harm to the environment or human health, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically." http://www.gdrc.org/u-gov/precaution-3.html.

This analysis using confidence intervals has classified this water body as 'healthy'. The ecological information collected by the FDEP indicates that eutrophication may be progressing even though the productivity response to nutrients is inconclusive. Therefore, the use of 90% confidence intervals for the geometric mean using all data (not just annual mean or median values) and threshold values determined from the productivity data (such as proposed by Briceno et al., 2010) as the numeric nutrient criteria appears to more protective for Florida's estuaries and coastal waters than the criteria currently proposed by EPA and is supported by the author.

References

Box, G.E.P., W.G. Hunter, and J.S. Hunter. 1978. Science for experimenters: an introduction to design, data analysis, and model building. Wiley Interscience, New York.

Briceno, H., J. Boyer, and P. Harlem. 2010. Proposed methodology for the assessment of numeric nutrient criteria for south Florida estuaries and coastal waters. National Park Service, Everglades National Park. Cooperative Agreement H5000-06-0104.

Florida Department of Environmental Protection, September 2010 Draft. Overview of approaches for nutrient criteria development in marine waters. Division of Environmental Assessment and Restoration, Standards and Assessment Section, Florida Department of Environmental Protection, Tallahassee, FL.

US Environmental Protection Agency. 2010. 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.