Development of Type III Site Specific Alternative Criteria for Nutrients

FDEP Bureau of Assessment and Restoration Support

October 24, 2011

Table of Contents

1. Purpose of Document

The purpose of this document is to provide information about data requirements and studies needed to support the establishment of Type III Site Specific Alternative Criteria (SSAC) for nutrients as described in Rule 62-302.800, Florida Administrative Code. The guidance provided in this document is also intended to ensure that data submitted to the Florida Department of Environmental Protection (DEP) are consistent with the requirements of DEP rules. The goal is that site-specific physical, biological, and water quality data are of suitable quality and sufficiency to evaluate the appropriateness of a Type III SSAC.

2. Type III Site Specific Alternative Criteria (SSAC)

2.1. SSAC Background

If the existing state-wide or regional water quality criterion is inaccurate for a given waterbody, a SSAC may be warranted, following the procedures outlined in Rule 62-302.800, F.A.C. The rule allows the establishment of three different types of SSAC when an affirmative demonstration is made that an alternative criterion is more appropriate for a specified portion of waters of the state. A Type I SSAC is based on natural background (minimally disturbed) conditions. For a Type II SSAC (which may include some human influences on water quality), it must be demonstrated that the criterion would fully maintain and protect the designated uses (human health and aquatic life), existing uses, and the level of water quality necessary to protect human health and existing and beneficial uses. For Type III SSACs, which are specific to nutrients, biological health assessments (evaluating both flora and fauna) are used to demonstrate full aquatic life use support, and the SSAC is established at levels representative of the existing associated nutrient regime (which is protective of the use).

In addition to SSAC type-specific requirements, all three types of SSACs must:

- Fully protect the designated use;
- If applicable, demonstrate support of the narrative nutrient criterion in subparagraph 62-302.530(47)(b);
- Be based on a sound, scientific rationale; and
- Protect downstream waters.

The Standards and Assessment Section recommends that any entity planning to petition for a SSAC meet with DEP staff before initiating studies in support of the petition.

2.2. Study Design for Type III SSACs

This section focuses on the scientific information essential for developing a Type III SSAC for nutrients. Type III SSACs require that biological health assessment data, collected in conjunction with nutrient data, demonstrate full support of healthy, well balanced aquatic communities [i.e., achieves the narrative nutrient criteria in paragraph 62-302.530(47)(b), F.A.C.]. Water chemistry, biological data (flora and fauna), and physical information from the waterbody are evaluated to determine if nutrient concentrations support well balanced communities of flora or fauna. Because of the complexity associated with nutrient enrichment effects, no single assessment tool is adequate to evaluate all potential impacts, and instead, a weight-of-evidence evaluation must be conducted.

For more information on use of biological assessment tools and a weight of evidence approach to evaluate whether or not floral and faunal imbalances are present in streams and rivers, see DEP-SOP-003/11, *Sampling and Use of the Stream Condition Index (SCI) for Assessing Flowing Waters: A Primer* (DEP-SAS-001/11), which includes:

- The nutrient enrichment conceptual model for streams;
- The process for numerically interpreting the narrative nutrient criterion in streams;
- Available procedures for evaluating the floral community in the stream, including chlorophyll *a* levels, periphyton abundance and species dominance (as measured using the Rapid Periphyton Survey [RPS]), and nuisance macrophyte distribution (as measured using the Linear Stream Vegetation Survey [LVS]);
- Evaluating the faunal community in the stream using the Stream Condition Index (SCI) or BioRecon;
- Efficiently collecting the information during one sampling event; and
- Examples of a weight-of-evidence approach for determining achievement of nutrient criteria.

In streams, if a site specific interpretation pursuant to paragraph 62-302.531(2)(a) (TMDL, SSAC, Level II WQBEL or RA Plan) has not been established, Nutrient Thresholds are used to interpret the narrative nutrient criterion in combination with biological information. The narrative nutrient criterion in paragraph 62-302.530(47)(b), F.A.C., shall be interpreted as being achieved in a stream segment if:

- Information on chlorophyll *a* levels, algal mats or blooms, nuisance macrophyte growth, and changes in algal species composition do not indicate an imbalance in flora or fauna; AND EITHER

- The average score of at least two temporally independent SCIs performed at representative locations and times is 40 or higher, with neither of the two most recent SCI scores less than 35 (*i.e.*, no faunal imbalances), OR

- The Nutrient Thresholds (expressed as annual geometric means) in the Table in section 62-302.531(2)(c) F.A.C., are not exceeded more than once in a three year period.

In cases where the Nutrient Thresholds are exceeded but there are no imbalances in both aquatic flora (phytoplankton, periphyton, vascular plants) AND fauna (invertebrate community), the narrative criterion is achieved. Sites with healthy flora and fauna are eligible for a Type III SSAC, provided that the loading of nutrients from the waterbody are limited as necessary to provide for the attainment and maintenance of water quality standards in downstream waters.

The number of stations required to be sampled is dependent upon the homogeneity of the stream. SSACs may be established for multiple stream segments that have homogeneous nutrient concentrations. SCIs must be conducted at least two spatially-independent stations in each homogeneous stream segment for which a SSAC is requested, and there must be at least two temporally independent SCIs conducted at each station. To demonstrate a healthy, well balanced faunal community, the average score of the SCIs must be 40 or higher, with neither of the two most recent SCI scores less than 35. SCIs collected at the same location less than three months apart are considered to be one sample, with the mean value used to represent the sampling period.

For lakes, at least two temporally independent Lake Vegetation Index (LVI) assessments must be conducted, with an average score of 43 or above. For both streams and lakes, the bioassessment data (SCI and LVI data) must be collected within the same years as the water quality data that is used to establish the SSAC, and for multiyear studies, at least one of the biological assessment must be conducted during the final year of the study.

The following elements should be addressed when designing a study in support of a Type III SSAC:

1. Biological sampling locations should be selected to reduce or eliminate the effects of confounding variables. Sampling should be conducted in areas where other physical factors, especially habitat and hydrology, do not limit biological expectations. Efforts should be taken to establish sites in stream reaches with minimal hydrological modifications and optimal habitat, including adequate substrate diversity and availability, intact stream morphology (minimal or no artificial channelization), adequate velocity and flow, and optimal riparian buffer zones (see DEP SOP FS 3000 for Habitat Assessment procedures). Sites should also be selected where light penetration through the tree canopy is representative of the stream segment (i.e., avoid bridge or powerline crossings where the canopy has been artificially reduced). Additional information on controlling for the effects of confounding factors is presented in the SCI and LVI Primers. If the entire stream reach is characterized by habitat and hydrological limitations, it is unlikely that a Type III SSAC is appropriate, and other options, such as conducting a Use Attainability Analysis, are available.

2. All SCIs or LVIs should be conducted consistent with the SCI and LVI standard operating procedures (SOPs) and Primers.

3. Sufficient water quality data (including TN, TP, and chlorophyll a) should be collected during representative conditions (*e.g.,* climatic, water level) and at locations that are suitable for linking the water quality to the biological data. This involves characterizing the water quality conditions, including temporal variability, that are associated with the biological data. For more information on data sufficiency and representativeness, see Section 2.3 below.
    a. The frequency and duration of the sampling needed is dependent upon the relative homogeneity of the system and the variability of the data. The DEP recommends quarterly sampling as a minimum sampling frequency and a minimum sampling duration of three years to adequately characterize annual variability.
    b. The DEP recommends a minimum of two water quality monitoring stations, but the specific number of water quality stations needed is dependent upon the size of the system, relative homogeneity of the system, and upon the availability of existing historical water data.
    c. Water quality stations must be located where there is a clear relationship between the nutrient regime and the system's biological health, as assessed using either the SCI or LVI. For streams, this generally means that the SCI sampling site should be downstream of or co-located with the water quality sampling station. For example, if a discharge or tributary significantly influences the nutrient concentrations in an area associated with the biological collection site, then data from stations located upstream of that discharge should not be used for establishing the SSAC values.
    d. Because a Type III SSAC relies on a distributional data analysis, which is sensitive to extreme events, data should not be collected during extreme climatic or hydrologic conditions, such as floods, droughts, or hurricanes (also discussed in Section 2.3 below).

4. Proper sampling station reconnaissance is vital to the overall sampling process. After conducting an initial desktop review using maps and aerial photographs, followed-up by a field reconnaissance of the site, the investigator should discuss the sampling locations with DEP Standards and Assessment Section (SAS) staff. Photos of the sites are part of the reconnaissance, and these photos, along with any additional information about the sites, must be available for SAS staff review. The use of existing stations is acceptable, but new stations may be required to ensure that sampling is representative of waterbody conditions.

5. To demonstrate that downstream waters are attaining water quality standards related to nutrient conditions, the petitioner should initially review available information on the DEP's website to see if any downstream waters have been placed on the verified list as impaired for nutrients, pursuant to Chapter 62-303, F.A.C. If the downstream waters attain water quality standards related to nutrient conditions, protection of downstream waters has been demonstrated. However, if the downstream waters do not attain water quality standards related to nutrient conditions, a demonstration must be made that the nutrient levels established by the Type III SSAC, when delivered to downstream waters, either:
    a. meet the allocations of a downstream TMDL; or

b. provide for the attainment and maintenance of water quality standards, using water quality models or other scientifically defensible methods.

2.3. Statistical Considerations for Developing Type III SSACs

The information presented in this document is intended to provide potential SSAC petitioners with guidance regarding the derivation of alternative numeric nutrient thresholds. It is not intended to provide a foundation in statistical methods or concepts, and a statistical text should be consulted for a thorough understanding of statistical concepts and techniques.

When developing numeric nutrient criteria (NNC) to protect and maintain a healthy, well-balanced community, it is important to account for natural variability in both the nutrient regime and in the biological communities, as well as other influences on the ecosystem. Derivation of nutrient criteria must be based on a sound scientific rationale, which requires adherence to the DEP's QA Rule (Chapter 62-160, F.A.C.) and identification of a reasonable ecological linkage between nutrients and protection of the designated use. The criteria should also account for and manage confounding factors during derivation, and control for Type I errors (incorrectly concluding that a system is impaired, when it is actually healthy [a "false positive"]) and Type II errors (incorrectly concluding that a system is healthy, when it is actually impaired [a "false negative"]). Statistical techniques should be selected to manage errors and explain variabilitiy.

The data sufficiency requirements needed to confidently determine a protective nutrient regime are dependent on the given environmental situation and the parameter's spatial and temporal variability. Criteria expression must account for natural fluctuations in the waterbody condition, and determining an appropriate nutrient regime will involve empirical evidence or water quality modeling. Monitoring must be of sufficient frequency to estimate the variability of the system so that the concepts of magnitude, duration, and frequency are properly accounted for when establishing protective nutrient criteria.

When deriving a water quality criterion, it is necessary to express the parameter concentration with respect to its magnitude, duration, and frequency. **Magnitude** is a measure of how much of a pollutant may be present in the water without an unacceptable adverse effect. **Duration** is a measure of how long a pollutant may be above the magnitude, and **frequency** relates to how often the magnitude may be exceeded without adverse effects.

For Type III SSACs, the magnitude shall be set at a level that maintains the current data distribution of a healthy existing condition, accounting for natural temporal variability. The magnitude component can be set to maintain the long-term central tendency (*e.g.*, geometric mean) of the distribution, while the frequency and duration components describe how often, and by how much, the nutrient concentrations can be above the central tendency while still being consistent with the baseline distribution.

Monitoring to support SSAC development must be of sufficient frequency to estimate the variance of the system, so that the concepts of magnitude, duration, and frequency are properly accounted for when establishing protective nutrient criteria. If monitoring frequency is not sufficient, it is likely that the derived criteria will not achieve the assumed Type I and Type II error rates; that is, incorrect future decisions regarding the nutrient impairment status will be made more often than assumed.

In terms of data sufficiency, the following should be taken under consideration when determining the number of samples needed:

- The quality of the data to be used;
- The spatial and temporal variability of the water quality constituent;
- Measurement errors associated with sampling and testing;
- The appropriateness of statistical treatment of the data and the rationale for its selection, including the handling of values less than the detection limit (generally, one half the detection limit is a good estimate if detection limits are consistent); and,
- That data were collected at suitable sites and during appropriate conditions to evaluate the parameter of concern.

Ideally, monitoring and assessment should be of sufficient rigor to detect significant relationships between anthropogenic nutrient inputs and biological responses (referred to as statistical power). Simultaneously, the assessment should not falsely indicate there is a human-induced effect when in fact there is none (low Type I error). Conversely, the assessment should minimize the rate at which true human-induced effects are not identified (Type II error).

It is important to note that statistical Type I and II errors are related to the null (H_o) and alternative (H_A) hypotheses and not whether the waterbody is achieving its designated use. In general form, the H_o states that the mean of (future) monitoring data is not greater than the baseline or reference long-term mean condition; while the H_A states that the mean of (future) monitoring data is greater than the baseline or reference long-term mean condition. Attainment of the designated use (and narrative nutrient criterion) is a related issue; however, decision errors related to the attainment of designated use are not strictly speaking Type I or II errors. In fact, these attainment decision errors occur because the wrong null hypothesis or baseline condition is being evaluated; that is, the criterion is either overly stringent or under-protective than is necessary. Statistical error rates may be assessed once proper null and alternative hypotheses are stated and an appropriate and representative baseline distribution has been established (**Table 1**). An appropriate baseline distribution is a nutrient data distribution that has been documented to be associated with maintenance of natural populations of flora and fauna (*e.g.*, passing SCI or LVI, absence of algal blooms or nuisance algal mats).

Table 1. Hypothesis testing decision error framework.

Decision	True Environmental Condition	
Decision Made by Applying Statistical Hypothesis Test to Data	Waterbody Achieves Reference or Baseline Condition	Waterbody Does Not Achieve Reference or Baseline Condition
Decide that Waterbody Achieves Reference or Baseline	Correct Decision	Decision Error (False Acceptance or type II error)
Decide that Waterbody Does Not Achieve Reference or Baseline	Decision Error (False Rejection or type I error)	Correct Decision

The ability to set accurate and scientifically defensible distributional based numeric thresholds requires an accurate characterization of the baseline or reference central tendency (*e.g.,* median, arithmetic average, geometric mean) and temporal and spatial variability (*e.g.,* variance, percentiles). In turn, the ability to accurately characterize the baseline distribution is dependent upon five interacting factors: sample size, variability, level of significance, power, and minimum detectable effect.

1. *Sample size*: Large sample sizes improve the accuracy of estimates of central tendency and variability and increase the ability to detect a difference between two groups of samples;

2. *Variability*: Variability is often expressed in terms of variance, percentiles, or range and typically reflects the spread of data around the central tendency. The more variable a parameter is, the higher the required sample size to accurately characterize the central tendency and variance. Additionally, increased variability reduces the ability to detect significant change from baseline unless corrected for by increased sample frequency;

3. *Level of significance*: This refers to the probability that an apparently significant difference is not real but simply due to random chance. This is referred to as alpha (α). An α of 0.10 means there is a 1 in 10 chance that an observed difference is due to random chance alone, or a test is 90% confident. The frequency at which the null hypothesis (H_o) is erroneously rejected is α. The erroneous rejection of the H_o is termed a Type I error (**Table 1**). A significance of α <0.1 is commonly used in environmental studies, but may be adjusted based on sample size;

4. *Power*: The probability of detecting a difference when in fact one exists; designated (1-β). β or a Type II error, is the probability of incorrectly concluding that two groups of samples are the same when in fact they are different. A β value less than or equal to 0.2 is commonly used in environmental studies. Decreasing the value of α (decreasing the probability of Type I errors) will increase the value of β (*i.e.,* increasing the probability of Type II errors), at a given sample size. Furthermore, variability influences such that the Power of a test is reduced as variability increases. Power can only be improved by increasing sample size; thus, the sampling frequency required to support a site specific evaluation (*e.g.*, SSAC, stressor identification) will be dependent on both the estimated variability of the parameter and the α value selected. The

variability of the parameter can be estimated from existing historic water quality data from the study waterbody or a nearby similar waterbody.

5. *Minimum detectable difference (MDD)*: Determining how much change is unacceptable should be linked to the inherent error associated with a given measurement system. The analytical detection limit is important during statistical comparisons, especially relating to the concept of a minimum detectable difference.

A SSAC should be derived with an objective of having no greater than a 10% expected Type I error rate to minimize erroneous conclusions that a water is impaired. DEP will consider lower Type I errors on a case by case basis in situations when the variance of nutrient is well quantified, such as a long data record (*e.g*., monthly for 10 to 20 years) or when an independent variable (*e.g*., color, salinity) can be used to explain a large portion of the variability in the nutrient parameter (**Figure 1**). An acceptable excursion frequency can be set using a three-year or five-year period as the basis of assessment. The exceedance frequency should account for inter-annual nutrient patterns and be established at a frequency that allows for effective and timely nutrient control; that is, it should account for and allow natural inter-annual variability associated with climatic cycles, and recognize that multiple high nutrient years can occur in succession. A consideration of this inter-annual correlation would suggest that the excursion frequency should allow for multiple excursions in a five-year period, such as two out of five or three out of five years.

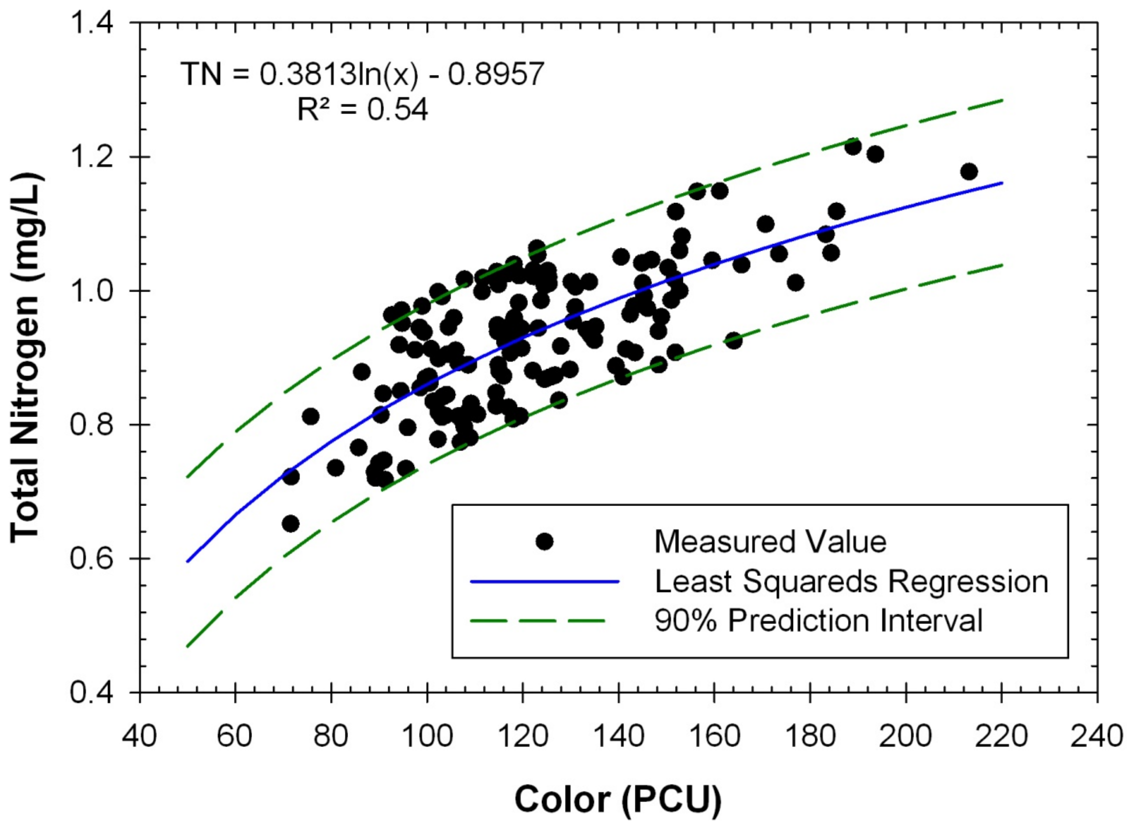

Figure 1. Relationship between color and total nitrogen in Boggy Creek. Color explains 54 percent of the variability in total nitrogen. A potential SSAC for TN in Boggy Creek could be established using the upper end of the 90% predication interval (equivalent to the one-side upper 95% prediction limit) and expressed such that no more than 5% of the future measurements shall be above the upper end of the 90% predication interval [i.e., TN = 0.3794ln(Color) - 0.7655]. The Boggy Creek TN SSAC would be expected to have a no greater than 5% Type I error rate.

Once an acceptable excursion frequency has been selected, a nutrient target should be set at a level that is expected to result in no more than a 10% Type I error rate, given the observed variability in the baseline dataset. The target is set at a percentile or upper prediction interval that corresponds with an x-year cumulative exceedance probability of no more than 0.9 or 0.95, summarized in Table 2 for durations ranging from three to five years. For example, an exceedance frequency of no more than once in a 5-year period should be set at the long-term 90th percentile (**Table 2**). Although DEP will consider alternative frequency and duration expressions for SSACs, DEP prefers consistency and thus recommends establishing alternative criteria at either the 80^{th} or 90^{th} percentile to be expressed as either an annual geometric mean not to be exceeded more than once in a three-year period or more than once in a five-year period, respectively.

Table 2. List of percentile targets required to achieve the specified α value for annual geometric mean concentration assessment periods ranging from three to five years and acceptable exceedance frequencies. For example, a SSAC expressed as not to be exceeded more than twice in a five-year period could be established at the 75th percentile and would be expected to have a Type I error rate of no greater than 10% (α =0.1).

Assessment Period (Years)	Exceedance Frequency	Percentile α =0.1	Percentile α =0.05
3	0	97	98
3	1	80	86
3	2	54	63
4	0	97	99
4	1	86	90
4	2	68	75
4	3	44	53
5	0	98	99
5	1	89	92
5	2	75	81
5	3	58	66
5	4	37	45

The primary objective of the existing condition approach is to establish magnitude and frequency limit(s), which if exceeded in the future, would allow one to conclude with sufficient statistical certainty that the new distribution is not consistent with the baseline distribution. In other words, DEP wants to be confident that future monitoring data are consistent with the baseline dataset distribution, rather than from some different data distribution. Given this goal, the use of a **"prediction interval"** is typically the most appropriate statistical tool for baseline data sets with at least four years in the period of record. Prediction intervals are used to estimate the range of future data, such that 100(1-α) % of the future data will fall within the prediction interval and 100(α) % will fall outside the interval (Helsel and Hirsch 2002). The upper 90 percent limit represents an estimate of the true long-term 90th percentile. Helsel and Hirsch (2002) provide an equation (**Equation 1**) to calculate an asymmetric (log-normal) prediction interval. An upper prediction limit is calculated as:

$$P.I. = e^{\left(\bar{y} + t_{(\alpha, n-1)} * \sqrt{\sigma_{\bar{y}}^2 + \sigma_{\bar{y}}^2 / n}\right)} \quad (1)$$

where,

= the mean of the log transformed data

$t_{(\alpha,n-1)}$= one-sided Students t statistic at n-1 degrees of freedom

σ^2_y = the variance of the log transformed data n=sample size (number of years)

The upper limit is used because the resulting value represents a level that should not be routinely exceeded, resulting in maintenance of current conditions or lower. In other words, if the prediction limit is not exceeded, there is confidence that nutrient concentration conditions have not increased. These limits correspond to annual geometric mean concentrations that are expected to be higher in only 10% of future years, given the range of spatial and temporal variability measured during the baseline periods for these waters. Therefore, it also represents a level that would be expected to result in a no more than 10% Type I error if applied as an annual geometric mean, not to be exceeded more than once in a either a 3- or 5-year period (see **Table 2** for appropriate percentile).

The prediction interval described above is prone to over-estimation of the true long-term 90th percentile, and thus increased Type II errors, when calculated based on a small sample size (*e.g.*, less than four years). For this reason, DEP does not recommend the use of the prediction interval for SSACs derived based on data sets spanning three years or less. In this case, DEP recommends calculating the criteria as 90th percentile calculated using non-parametric methods (*i.e.*, ranking the data) or assuming that the data follow a standard normal cumulative distribution. Because nutrient and chlorophyll are typically skewed to the right (*i.e.*, approximate a log-normal distribution), it is usually advisable to log (natural log) transform the data prior to calculating the percentile based on the standard normal cumulative distribution.

The 90th percentile of annual geometric means, assuming a log-normal distribution, is calculated as

(2)

where 1.2816 (substitute 0.6745 to compute a 75th percentile) is the inverse of the standard normal cumulative distribution with a probability of 0.9, is the mean of natural log transformed annual geometric means, and σ is the standard deviation of the natural log transformed annual geometric means (**Equation 2**). Statistical and spreadsheet computer programs typically include functions that can be used to calculate the expected percentile based on an assumed distribution (*e.g.*, NORMINV, LOGINV).

Both the prediction interval and the log-normal distribution estimated percentile (**Equation 2**) are parametric statistical techniques, which are based on an assumption that the data are from a normal distribution, or one that can be made normal through a transformation. Estimation error may occur if the data distribution significantly deviates from a log-normal or normal distribution. Petitioners should investigate the data distribution; however, the assumption of log-normality can only be verified with large datasets, such as those with over 200 data points. It is acceptable to assume a log-normal distribution even if deviations from a true log-normal distribution occur at the tails of the sampled distribution (*i.e.*, ≤ 5th or ≥ 95th percentile), as long as the fit is very good at the upper percentile under consideration (*e.g.*, 75th or 90th).

If the data distribution deviates significantly from either a log-normal or normal distribution, DEP recommends using non-parametric statistics such as a 90th percentile of the data based on ranks to derive a SSAC. However, a percentile calculated in this way only reflects the upper distribution of the baseline data set and makes no attempt to account for uncertainty; therefore, DEP advises that petitioners collect additional data to better characterize the data distribution particularly in cases with fewer than 30 to 50 data points. Alternatively, in cases with greater than 100 data points, petitioners may use bootstrapping techniques to estimate a 90% confidence interval around a 90th percentile and use the upper end of this confidence interval as the numeric expression of the SSAC, which will account for uncertainty.

The statistical methods described in the previous paragraphs are approaches that DEP has used in the past to derive nutrient thresholds and are consistent with the methods used to develop the Nutrient Watershed Region Nutrient Thresholds in Rule 62-302.531, F.A.C. However, an entity has the option of petitioning for a SSAC derived using alternative statistical methods as long as the petition describes the statistical assumptions as well as how the proposed threshold is consistent with a Type I error rate of no greater than 10%.

While Section 2.2 provides an example data period of record of three years, a petitioner has the option of collecting more data or utilizing previous (found) monitoring results. The petition must demonstrate that the waterbody achieved the narrative nutrient criteria in paragraph 62-302.530(47)(b), F.A.C., throughout the period of record used to establish the SSAC value. This demonstration is made by collecting Biological Health Assessments both at the beginning and end of the SSAC study. The biological health of the waterbody preceding the SSAC study can be demonstrated using previously collected Biological Health Assessment data obtained from DEP or other sources. Alternatively, the petitioner could demonstrate that there has not been a statistically significant trend in nutrients throughout the expanded period of record including the study period. This alternative is based on the logical argument that if concentrations during the SSAC study are protective of healthy biology and nutrient concentrations have not changed over the period of record, then the historic concentrations must have been protective of healthy biology. However, if nutrient concentrations have decreased and there are no biological data to demonstrate that the previous high nutrient levels were associated with healthy biology, it cannot be assumed that the waterbody supported healthy biology during the entire period of record. Consequently, the SSAC must be derived using only data from the study period.

For example, a petitioner wishes to pursue a Total Phosphorus SSAC for Clear Creek in the Peninsula Nutrient Watershed region. Clear Creek has been monitored on a roughly quarterly basis by the county since January of 2000, but only had one SCI collected by DEP in 2006. The petitioner initiated a three year SSAC study in 2010 and collected two additional SCIs at two spatially independent stations. The data collected during the study were combined with county data to provide a thirteen year period of record (**Figure 2**). Note: there were sufficient water chemistry and Biological Health Assessment data following the second SCI collected in May 2010; however, the petitioner wanted to collect two additional of years of data to verify the county data and to better characterize temporal variability through increased monitoring frequency. The five passing SCI scores clearly demonstrate that the stream supported healthy biology between October 2005 and the end of 2012. Additionally, a Mann's trend test demonstrated that there was not a statistically significant trend across the entire period of record; therefore, DEP agreed that it was appropriate to use the entire period of record to derive a protective TP SSAC (**Table 3**) using the prediction interval approach described in **Equation 1**.

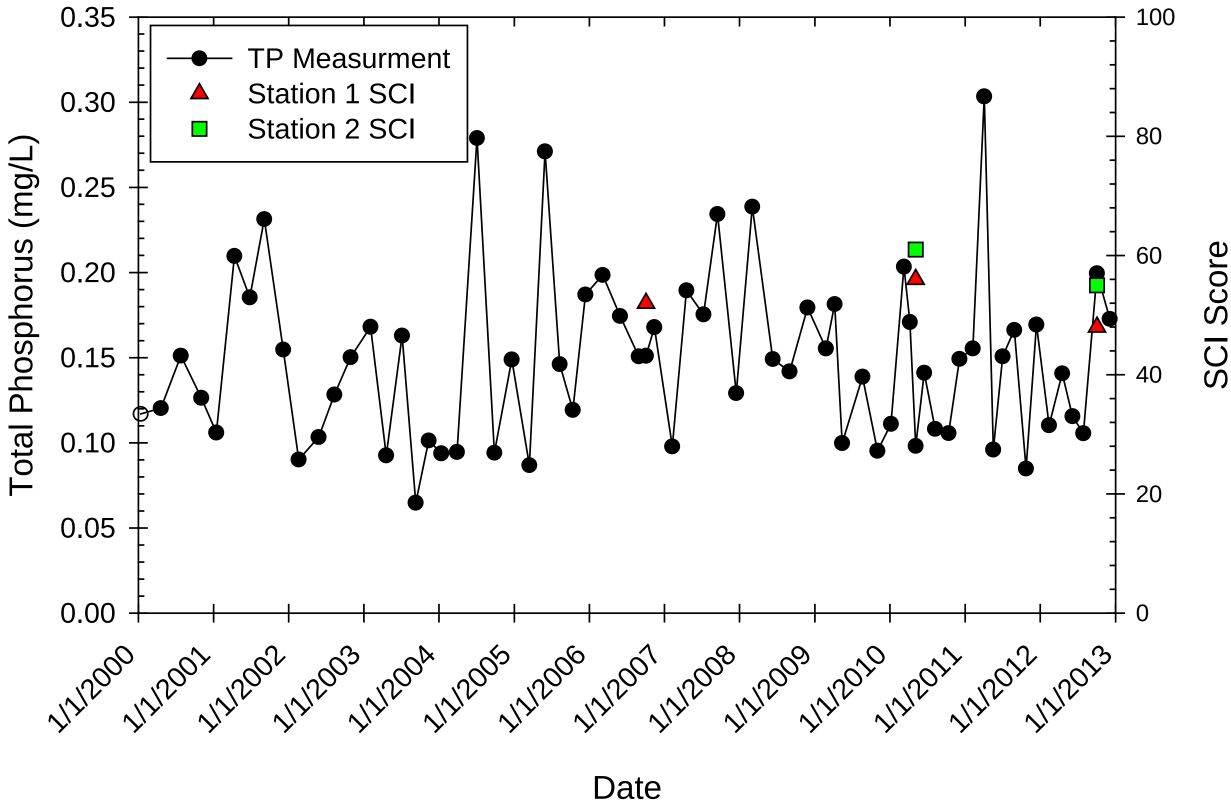

Figure 2. Time series of total phosphorus concentrations and SCI scores from Clear Creek. SCI scores were used to demonstrate healthy biological condition and total phosphorus measurements were used to derive a SSAC for the creek.

Table 3. Summary of annual total phosphorus concentrations in Clear Creek and derivation of the SSAC threshold using a 90% upper prediction limit. The SSAC is expressed as a value not to be exceeded more than once in a five-year period.

Year	Mean Ln TP	Geometric Mean TP	N
2000	-2.055	0.128	4
2001	-1.764	0.171	5
2002	-2.156	0.116	4
2003	-2.201	0.111	5
2004	-2.054	0.128	5
2005	-1.895	0.150	5
2006	-1.786	0.168	5
2007	-1.846	0.158	5
2008	-1.752	0.173	4
2009	-2.040	0.130	5
2010	-2.026	0.132	8
2011	-1.904	0.149	7
2012	-1.990	0.137	6
Period of Record Mean	-1.959		
Standard Deviation	0.152953		
t	1.36343		
Prediction Limit	0.176		

Conversely, if data and statistical trend analysis indicate that there has been an increasing nutrient trend over time and the study demonstrates healthy biology throughout the SSAC study period, then the petitioner may either derive a SSAC threshold using only data from the recent study period or may detrend the historic nutrient data to a level consistent with the study period average. Detrending in this manner would allow the petitioner to use the entire period of record to better characterize the long-term inter-annual variability. The increasing nutrient trend must be found to be true temporal change and not due to inconsistencies between laboratories, study sites or other methodological differences. These types of differences will typically occur as step changes rather than continuous trends.

3. Applicability of Quality Assurance (QA) to Type III SSACs

Consistent with the QA Rule, DEP's Quality Assurance Directive states that it is DEP's policy to use scientifically valid and legally defensible data for protection of the environment. DEP must ensure that all scientific work products and environmental decisions are supported by sound science. Performing the necessary QA steps will ultimately save time and staff resources and result in a correct DEP action.

All personnel scheduled to conduct bioassessments (SCI, BioRecon, LVI) must complete at least eight hours of DEP sanctioned field training and pass field audits for the bioassessment to be performed. Biological field evaluations have the potential to contribute an enormous amount of information to an overall water body assessment, but it is critical that staff follow all SOPs, particularly those dealing with proper conditions for the bioassessment, as required by DEP-SOP-003/11 SCI 1000 and *Sampling and Use of the Stream Condition Index (SCI) for Assessing Flowing Waters: A Primer* (DEP/SAS/001/11) and DEP-SOP-003/11 LVI 1000 and *Sampling and Use of the Lake Vegetation Index (LVI) for Assessing Lake Plant Communities in Florida: A Primer* (DEP/SAS/002/11) available at: http://www.dep.state.fl.us/water/bioassess/training.htm).

There are several ways to ensure that data are appropriate for numeric nutrient study objectives:

- **Do the data make sense?** When using historical data collected by other monitoring entities, it is important to ensure that the data have been collected for purposes that are consistent with numeric nutrient study objectives. For example, the location and hydrologic conditions under which the data were collected may not represent the ambient conditions of a waterbody (*e.g.,* samples were from a treatment system, not ambient waters). Deciding if data are appropriate for a given use is governed by the QA Rule and the document titled, "Process for Assessing Data Usability" (DEP-EA 001/07, found at: http://www.dep.state.fl.us/water/sas/qa/index.htm). The investigator must use knowledge of the area (and tools such as the Landscape Development Intensity Index and Geographic Information Systems) to determine if any previously collected data involve unrealistic results, such as if there are undisturbed sites that have unexplained values. Data must be carefully evaluated to assure that they are usable for evaluating attainment of the nutrient criteria. The DEP Standards and Assessment Section should be consulted to ensure that the final study conclusions are supported by appropriate data of suitable quality.

- **Was the sampling site suitable**? The sampling site location is critical in determining what the water quality data actually represent. Sampling locations should represent the ambient conditions of the waterbody. First, the investigator must make sure that the proper system type is being evaluated (*e.g.,* wetland systems have different biological and nutrient expectations than do streams or lakes, so the system type must be correctly identified). Also, channelization near road crossings, obstructions (improperly placed culverts), isolated clear cutting (unusually open canopy cover), can influence water quality results within the disturbed localized area, and results from such areas may not represent the waterbody reach. The focus of nutrient studies is to determine if anthropogenic nutrient loading is causing adverse biological effects, so data must be obtained from stations that provide information relative to that objective. For example, samples collected during rain events at a stormwater outfall are not representative of the waterbody reach, where assimilative capacity and dilution would influence the results found in the waterbody proper.

  If physical degradation (poor habitat and hydrologic modification) in a particular stream segment was associated with biological health assessment failures (*e.g.,* average of the two most recent temporally independent SCI <40), the biology should be re-sampled at an area where there is adequate habitat and less hydrologic modification. If the average of two SCIs is 40 or higher and there are no floral imbalances (see weight of evidence approach described in the SCI Primer) in a stream segment with adequate habitat and an intact hydrologic regime, this indicates that water quality (nutrients) is not a limiting factor.

- **Were hydrologic conditions appropriate?**

  Hydrologic conditions can significantly influence both biological and water quality results. Throughout Florida, there may be extreme differences between water levels and flow conditions between periods of low rainfall and periods of high rainfall. Water quality data used to assess establish SSACs should be collected over a period of time that is representative of the typical hydrologic conditions. Thus, DEP recommends a minimum of three years of data be evaluated for any of these purposes, although longer time periods of five to ten years are preferred and will better characterize long-term variability and control statistical errors.

  During droughts, streams may be disconnected pools (or totally dry), lakes may have their littoral zones exposed (organic sediments may be oxidized), and wetlands will be dry (again, oxidizing sediments). During such conditions, leaf litter/plant material decomposition may primarily be responsible for naturally low DO, soil/sediment leaching may result in higher nutrient concentrations, and chlorophyll-*a* may increase due to the stagnant conditions. Conversely, hurricanes or other flooding events can mobilize soil organic and inorganic materials (metals, nutrients) from floodplains, and create a situation where no colonized biological habitats are available (SCI sampling not possible). High organic nutrients and low DO have been shown to be a natural occurrence in streams and estuaries during and after hurricanes.

  Unless a large data set is involved (*e.g.,* monthly samples for 20 years), these extreme hydrologic events may not be representative of typical ambient conditions and water quality conditions during these events may overly skew the data distribution. These events are representative of the full range of natural variability, but may be overly influential when evaluating shorter periods of record. DEP recommends that water quality data collected during extreme hydrologic conditions, such as flood and drought events that recur less than once in a twenty-five year period, should be evaluated for potential exclusion. If the data set is sufficiently large (*e.g*., monthly for 10 to 20 years), these events are less likely to be overly influential and could be included because they are representative of the full range of natural variability. The effect of extreme event data on the overall distribution (*i.e.,* on the 90^{th} and 95^{th} percentiles) should be evaluated and overly influential data should be excluded if it can be demonstrated and documented that these data were associated with unusual hydrologic conditions.

  DEP will require that all excluded data be identified and clear documentation as to the basis for exclusion must be provided in the supporting documentation. Furthermore, any SSAC would need to clearly stipulate that future data collected under similar extreme conditions be excluded from attainment assessments.

- **Were DEP SOPs followed?**
  Documentation to affirmatively support that the sample integrity was maintained throughout the process is required by the QA Rule, and must be made available for DEP staff review. This evidence includes sampling procedures, equipment and container suitability, preservation, holding times, and other items required by the QA Rule. DEP SOPs are found at: http://www.dep.state.fl.us/water/sas/sop/index.htm

- **Was the lab performing the analyses certified by the National Environmental Laboratory Accreditation Program (NELAP)?**
  NELAP certification status may easily be checked at:

  http://www.dep.state.fl.us/labs/cgi-bin/aams/index.asp

- **What did data qualifiers indicate and were the data usable?**
  Qualifiers and other information should be evaluated to determine if the data were suitable for the intended purpose, following the "Process for Assessing Data Usability," found at: http://publicfiles.dep.state.fl.us/dear/sas/sopdoc/2008sops/usability_doc.pdf

All of these QA issues must be considered, with emphasis on sample representativeness and data defensibility. Incorrect or inappropriate data should not be used for interpreting the narrative nutrient criteria as described in this document; and instead, valid data to support monitoring objectives should be collected. The key issue is whether the samples accurately represent the study objectives for developing a Type III SSAC.

4. Example Type III SSACs

4.1.TP SSAC for portions of the Alafia River

The Mainstem of the Alafia River (WBIDs 1621A, 1621B, and 1621C), located in Hillsborough County within the West Central Nutrient Watershed Region, drains into Hillsborough Bay (**Figures 3 and 4**). The stream cannot be characterized as minimally disturbed. Approximately 36 percent of the watershed has been mined (extractive land use). Only 17 percent of the basin is residential, with the majority of the residential areas around the lower estuarine reach, and the remainder in the upper reaches of the North Prong watershed near Lakeland. Facilities that discharge to surface waters are required to be permitted through the National Pollutant Discharge Elimination System (NPDES). There are 23 active NPDES permitted discharges in the Alafia River watershed.

Figure 3. Photos of the Alafia River near the lower portions of WBID 1621A. Note the tidal influence, meaning SCI is a not an appropriate tool in these areas.

Figure 4. Station ALAFTP600, Alafia River at TECO powerline access gate. This is a freshwater site appropriate for SCI sampling.

There were five SCIs (with concurrent habitat assessment) conducted from the Mainstem of the Alafia between 2009 and 2010. Individual SCI scores have ranged from 38 to 52, with an average score of 47. Note that no score was below 35. The variability in SCI scores, even at minimally disturbed sites, may be explained by random, natural events such as sporadic, unpredictable rain and drought, which in turn influence the relative abundance of inundated substrates available for invertebrate colonization. These natural stressors (*e.g*., flood, drought, natural low substrate diversity, natural episodic low DO, etc.) will affect all sites, even those with minimal disturbance from humans. To determine when one particular water quality variable is responsible for adverse effects (causing an impaired or imbalanced community) one must reasonably account for and control these other factors. In 2010, two temporally independent Rapid Periphyton Surveys (RPS) were conducted at five sites in the Alafia River, including sites in the Mainstem, as well as the North and South Prongs. The RPS results indicate that a high percentage of sampling points for all five sites (ranging from 98.3% to 41.3%) had no algae. Filamentous algae of thickness > 6 mm (ranks 4-6) were infrequently observed (12-15%) in the Mainstem of the Alafia. Previous data show that these levels of algae are within the range of reference site conditions and that these algal levels do not interfere with a healthy biological community. Additionally, Linear Vegetation Survey (LVS) results showed that the vascular plant community was similar to reference streams. Taken as a weight of evidence, the biological data provides strong empirical evidence demonstrating that the Alafia River is fully meeting aquatic life use support expectations at the existing TP and TN regime.

Despite exhibiting a healthy biological condition, the Mainstem of the Alafia River exceeds the West Central Stream Nutrient Threshold one in three year annual geometric mean limit of 0.49 mg/L TP. Because the river has been demonstrated to be biologically healthy, the existing levels of TP are protective of the designated use, and a Type III SSAC for TP is justified. Furthermore, the TP SSAC would be protective of downstream waters (Hillsborough Bay) because: 1) algal growth limiting nutrient bioassay experiments clearly demonstrate that Hillsborough Bay is primarily and consistently nitrogen limited; 2) recent USGS publications demonstrate that TP concentrations in the Bone Valley stream sediments are extraordinarily elevated and phosphorus has always been the excess nutrient in the Alafia River even absent human activities; and 3) the historical record clearly shows that the TP concentrations in the Alafia River during the past five years (the basis for the TP SSAC) are associated with TP levels in Hillsborough Bay that comply with the protective 0.45 mg/L criterion proposed by the Estuary Program for full protection of seagrasses (**Figure 5**).

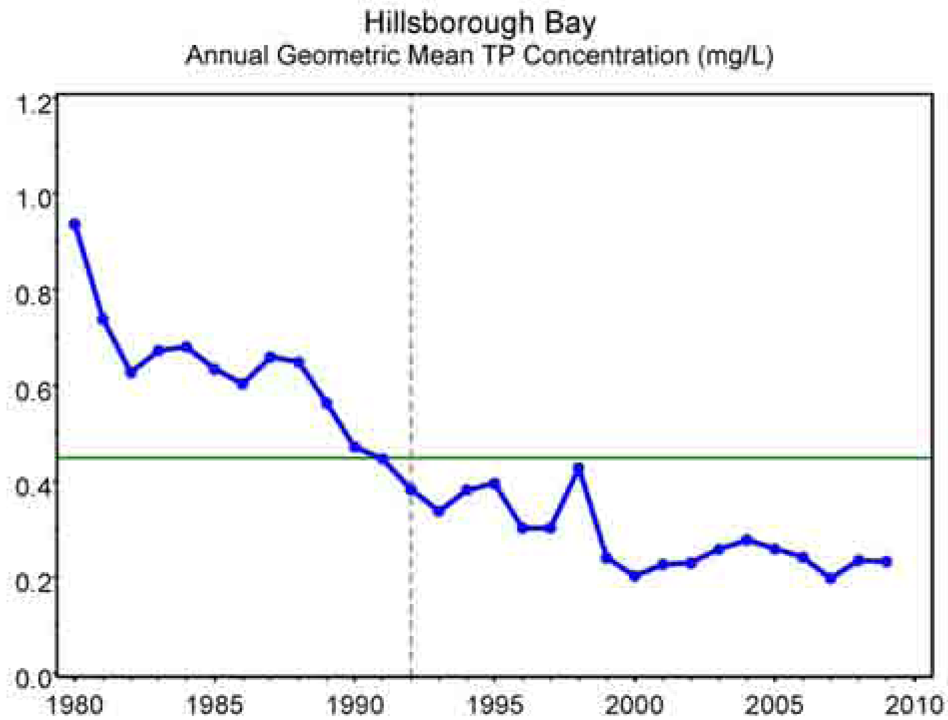

Figure 5. Annual geometric mean TP concentrations in Hillsborough Bay (receiving waters for the Alafia River) from 1980 to 2010 (from Janicki 2011).

DEP staff developed a protective TP SSAC using available TP data from STORET for the period from January 2006 through March 2010. This date range was selected because:

- This was the time period associated with the healthy Stream Condition Index and floral data;
- This period represented conditions when downstream chlorophyll-*a* targets in Hillsborough Bay were achieved; and
- There was sufficient data density to calculate a robust distribution.

Annual geometric mean total phosphorus levels were calculated using all monitoring stations in the Mainstem of the Alafia River and are summarized in **Table 4**. The TP SSAC limit was calculated based on the annual average log-transformed TP concentrations using an upper 90th prediction limit [equation (1)]. The SSAC calculation is summarized in **Table 4**.

Table 4. Annual geometric mean TP in the Mainstem of the Alafia River and derivation of the one-sided upper 90% Prediction Limit.

Year	Mean LN(TP)	Geometric Mean (mg/L)	Number of Samples
2006	0.014	1.01	84
2007	-0.018	0.98	91
2008	-0.014	0.99	72
2009	0.030	1.03	37
2010	0.158	1.17	26
Mean LN TP	0.034		
Period of Record Geometric Mean	1.03		
Std. Dev. TP	0.072		
n (years)	5		
$t_{0.1(1),n-1}$	1.533		
Prediction Limit (TP SSAC)	**1.17**		

Based on the biological results for the Alafia River system and a variety of evidence for Hillsborough and Tampa Bay, DEP has determined that the existing levels (past 5 years) of TP are fully protective of the Alafia River and the downstream receiving waters of Hillsborough and Tampa Bay. DEP conducted an analysis of the long term nutrient data and derived a TP SSAC based on the upper 90 percent prediction limit of the annual geometric mean levels. The TP SSAC is expressed as an annual geometric mean total phosphorus limits of **1.17** mg/L not to be exceeded more than once in a 5-year period. The SSAC is expected to result in no greater than a 10% type I statistical error rate consistent with Section 2.3 of this document (see **Table 2**).

4.2. Example of Type III TN SSAC for the Econfina River

The Econfina River (WBID 3402) in Taylor County is a minimally disturbed blackwater stream draining into the Gulf of Mexico (**Figure 6**). There are no point source discharges in the system and the watershed Landscape Development Intensity Index (LDI) is 1.47, indicating benign land uses in the basin. From 1996 to 1998, Stream Condition Index (SCI) sampling was conducted 18 times in the WBID as part of a study to evaluate the effectiveness of forestry best management practices. All 18 of these SCIs scored in the category referred to as “excellent” in the earlier version of the index, indicating that this site has a long demonstrated history of healthy biota. From 2004 to 2008, eight additional SCIs were performed in this WBID. These SCI scores ranged from 40 to 59, with an average of 50, demonstrating continued healthy biota. Chlorophyll *a* is low in the river and there is no evidence of excess algal mats or nuisance aquatic plant growth. The downstream estuary has a healthy Submerged Aquatic Vegetation community.

Despite exhibiting excellent biological health, the Econfina River exceeds the Panhandle East Stream Nutrient Threshold one in three year annual geometric mean limit of 1.03 mg/LTN. The DEP concluded that decomposition of leaf litter from extensive forested swamps is the source of this nitrogen. Because the river and downstream waters have been demonstrated to be biologically healthy, the existing levels of TN are protective of the designated use, and a Type III SSAC for TN is justified.

Figure 6. Photo of the Econfina River (Taylor County) near the transition zone to the estuarine area, within the Econfina River State Park.

From 1990 to 2010, a total of 325 TN results were available for this WBID. During this period, annual geometric mean TN ranged from 0.5 mg/L to 1.5 mg/L. Based on the existing data, a petitioner derived a TN SSAC by using the binomial distribution and establishing a value that would have no more than one exceedance in any 3-year period, based on a 5% Type I error rate (equivalent to the 86^{th} percentile of the data set). Using the binomial distribution with no more than one exceedance in any 3-year period, the percentile of the nutrient annual geometric mean distribution that is consistent with a type I error of 5% was selected. A Type I error of 5% was considered appropriate because of the robust data indicating the Econfina River and downstream estuary are healthy.

The TN data were shown to be log-normally distributed using the Kolmogorov-Smirnov test for a significant fit at 5%. Accordingly, all calculations were performed using log transformed nutrient data and the corresponding statistics. To appropriately account for inherent uncertainty in the nutrient statistical descriptors and therefore the computed upper percentile, a one-sided 90% upper confidence interval limit (UCL) was computed around the previously computed 86th percentile. The procedure is documented in "Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Unified

Guidance", March 2009 (EPA 530/R-09-007). The computed long term geometric mean TN value was 0.80 mg/L and the 86th percentile of the TN annual geometric mean distribution was 1.26 mg/L. The 90% UCL for TN around the 86^{th} percentile was calculated to be 1.49 mg/L. Therefore, 1.49 mg/L is the TN magnitude for the Type III SSAC, which shall not be exceeded more than once in a three year period. Although the statistical approaches used to derive the Econfina TN SSAC were different from those described in subsection 2.3 of this document, DEP reviewed the technical merit of the alternative approaches and found that data met the necessary statistical assumptions and the criteria would be consistent with a Type I error rate of no greater than 10%. Furthermore, DEP concluded that the proposed SSAC limits would fully support the designated use.