Establishing Nitrogen Endpoints for Three Long Island Sound Watershed Groupings:

Embayments, Large Riverine Systems, and Western Long Island Sound Open Water

Subtasks F/G. Draft Summary of Empirical Modeling and Nitrogen Endpoints

Submitted to:

U.S. Environmental Protection Agency
Region 1 and Long Island Sound Office

Submitted by:

Tetra Tech, Inc.

April 13, 2018

This Tetra Tech technical study was commissioned by the United States Environmental Protection Agency (EPA) to synthesize and analyze water quality data to assess nitrogen-related water quality conditions in Long Island Sound and its embayments, based on the best scientific information reasonably available. This study is neither a proposed Total Maximum Daily Load (TMDL), nor proposed water quality criteria, nor recommended criteria. The study is not a regulation, is not guidance, and cannot impose legally binding requirements on EPA, States, Tribes, or the regulated community. The technical study might not apply to a particular situation or circumstance, but is intended as a source of relevant information to be used by water quality managers, at their discretion, in developing nitrogen reduction strategies.

Subtasks F and G. Summary of Empirical Modeling and Nitrogen Endpoints

EPA and Tetra Tech developed a draft version of this memo dated November 15, 2017, which was circulated to the Technical Stakeholder Group (TSG). Appendix F1 addresses the technical comments submitted by members of the TSG on the draft memo. Appendix F2 presents a compilation of the comments submitted by the TSG, including letters submitted by Connecticut, Massachusetts, New Hampshire, Vermont, Rhode Island, New York State, New York City, and Suffolk County.

Contents

Subtask F. Summary of Empirical Modeling

Introduction

EPA selected seagrasses and other aquatic life for developing nitrogen endpoints. As described in the *Literature Review Memo*, these assessment endpoints are principally reflected by water column chlorophyll *a* (through its effect on light for seagrass growth) and dissolved oxygen (DO) (through its effect on benthic fauna and fishes). Support for seagrasses also implicitly supports a variety of other aquatic assessment endpoints including benthic fauna. Benthic fauna were not explicitly reviewed in the *Literature Review Memo*; this was largely a function of a lack of sufficient data and indicator tools for this group across LIS. However, benthic fauna are implicitly a valued aquatic life use endpoint associated with seagrass and DO protection, and their consideration was implicit in the selection of the endpoints for which there was abundant data across LIS (i.e., chlorophyll *a*, clarity, and DO). The value of using multiple endpoints is that for those habitats where seagrass may not be expected (e.g., deeper waters), DO considers protection of other aquatic life.

For this task, Tetra Tech conducted empirical analyses to support development of nitrogen endpoints for the selected watersheds (Figure F-1). Three complementary empirical approaches (lines of evidence) were used to identify candidate total nitrogen (TN) endpoints:

- Scientific Literature Analysis: Acquired literature-based nitrogen endpoints (loads and concentrations) associated with protection of comparable assessment endpoints (seagrasses and other aquatic life) in similar estuaries during development of the *Literature Review Memo*.
- Stressor-Response Analysis: Developed nitrogen endpoints from empirical statistical models of relationships between chlorophyll *a* and TN using LIS surface water quality data (*Subtask D: Summary of Existing Water Quality Data*). Identified chlorophyll *a* endpoints from additional empirical statistical models of assessment endpoints (seagrasses and aquatic life) for light availability (Secchi or light attenuation) and DO as a function of chlorophyll *a*. Modeled literature-based chlorophyll endpoints were also identified and used.
- Distribution-Based Approach:[1] Developed nitrogen endpoint concentrations from 25th percentiles of TN concentration distributions in embayment and open water stations.

From these endpoints, waterbody-specific nitrogen endpoints are identified in Subtask G.

[1] The distribution-based approach, often referred to as the reference-based approach, has been frequently applied for deriving endpoints for various applications, including TMDLs (USEPA 1999) and criteria (USEPA 2001).

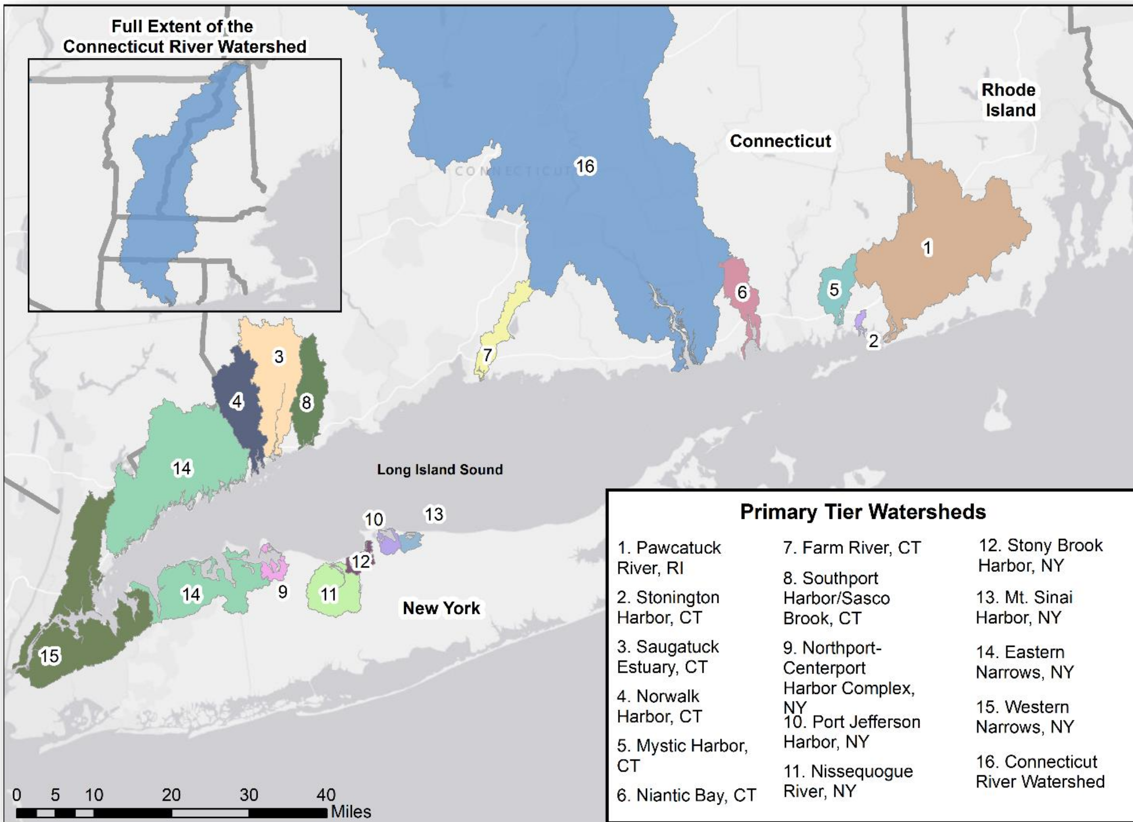

Figure F-1. Map of Study Area Indicating LIS Watersheds Evaluated

Methods

Scientific Literature Analysis

Tetra Tech conducted a literature review to evaluate biological, chemical, and physical aspects of possible assessment endpoints (seagrass, macroalgae, DO, phytoplankton, harmful algal blooms, and oysters) to protect designated uses. In conducting the search, Tetra Tech focused geographically on estuaries and bays from the Chesapeake Bay north to the state of Maine, and studies published between 1980 and the present. From the literature search, Tetra Tech identified literature-based TN values relevant to protection of designated uses in LIS. Tetra Tech also identified literature-based light/clarity values relevant to protection of seagrasses and summarized existing DO criteria relevant to protection of aquatic life in LIS. Additional detail is available in the *Literature Review Memo*. A decision was made to focus primarily on values from the most proximate study areas (Massachusetts) and not to incorporate values from farther north (Great Bay, NH) or south (Chesapeake Bay) because those systems were considered substantially different; the northern systems being farther from the Virginian province and the southern being a substantially different estuarine system in terms of size, geography, hydrodynamics, salinity structure, and climate. The approach assumes that the Massachusetts estuaries literature based targets were appropriate for LIS, given the similarities in geography, climate, and species composition (e.g., *Zostera marina*) consistent with similar physical and chemical habitat requirements in both embayment as well as shallow and deeper open water habitats between the two regions.

A synthesis of the resultant TN values is shown in Table F-1.

For embayments, Tetra Tech selected a median value of 0.40 mg/L TN to protect the seagrasses in embayments. This value is the rounded value of the median TN protective of seagrasses (0.39 mg/L; range: 0.30 to 0.49 mg/L). Values above the literature review maximum TN concentration of 0.49 mg/L were not considered protective of eelgrass (see Table F-1).

Tetra Tech identified a TN concentration of 0.80 mg/L as a severe degradation endpoint for embayments regionally. Generally, at TN concentrations at or above this level: macroalgal accumulation, near or complete loss of dissolved oxygen (DO), and fish kills will be observed; and benthic communities are often nearly absent during the warmer months or composed of only a few species of the most stress tolerant species; and seagrasses are generally absent. (Howes et al. 2003). Therefore, values at or above 0.80 mg/L were excluded from consideration leaving 0.60 mg/L as the maximum (see Table F-1). As a result, for open water segments, Tetra Tech used the median of all assessment endpoints (0.41 mg/L TN rounded to 0.40 mg/L) and range (0.30–0.60 mg/L TN).

Table F-1. Summary of Literature Review-based Endpoint Values for Total Nitrogen

	TN (mg/L)	Assessment Endpoint	Location	Citation
	0.49	Seagrass transplant survival > 50%	SE Massachusetts Estuaries	Benson et al. 2013[a]
	0.39	Seagrass transplant survival > 75%		
	0.42	Healthy seagrass		
	0.34	Seagrass survival		
	0.31	Restoration of eelgrass	Massachusetts Estuaries	MEP 2017[b,d]
	0.49	Restoration of eelgrass		
	0.30	Eelgrass present	SE Massachusetts Embayments	Howes et al. 2003[c]
	0.39	Eelgrass present		
Median	0.39	Summary for Seagrass Protection Endpoints (Used for Literature Line of Evidence for Embayments, N=8)		
Min	0.30			
Max	0.49			
	0.40	Infaunal habitat protection	Massachusetts Estuaries	MEP 2017[d]
	0.60	Infaunal habitat protection		
	0.41	Benthic habitat protection		
	0.91	Benthic habitat protection		
	0.50	Upper end of good/fair conditions and lower end of moderate impairment	SE Massachusetts Embayments	Howes et al. 2003[c]
	0.80	Severe ecological degradation begins		
	0.30	No macroalgae		
	0.50	Macroalgae might occur in some regions		
	0.39	DO generally >5 mg/L		
	0.50	DO generally >5 mg/L		
Median	0.41	Summary for All Endpoints		

	TN (mg/L)	Assessment Endpoint	Location	Citation
Min	0.30	(Values at or above the severe degradation endpoint of 0.80 were excluded, leaving a maximum of 0.6 – see narrative above; N=16) (Used for Literature Line of Evidence for Open Waters)		
Max	0.60			
Median	0.46	Summary for Non-Seagrass Endpoints (Values at or above the severe degradation endpoint of 0.80 were excluded, leaving a maximum of 0.6 – see narrative above; N=8)		
Min	0.30			
Max	0.60			

[a] Long term tidally averaged value; [b] Long term average; [c] Long-term, ebb tide average
[d] See Appendix F3 for additional information about the 33 documents included in this citation

Stressor-Response Analysis

For the stressor-response method, Tetra Tech developed empirical statistical stressor-response relationship models using surface water quality data described in *Subtask D: Summary of Existing Water Quality Data.* EPA identified three watershed groupings: embayments, riverine, and open water. For the stressor-response method, hierarchical regression was used to model embayment and riverine data together (the Connecticut River area of influence identified in *Subtask E: Summary of Hydrodynamic Analysis* was included as an additional "embayment"), and multiple linear regression was used to model open water observations.

The stressor-response models often produced values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

Hierarchical Models

Hierarchical models are also known as multilevel models, mixed models, and mixed effects models. These models are "mixed" in that they contain both global (so-called "fixed") parameters and group-adjusted (so-called "random") parameters. Hierarchical models can be viewed as an extension of linear regression models. A linear regression model of chlorophyll *a* as a function of TN takes the following form:

$$Chlorophyll_i = \beta_0 + \beta_1 * TN_i + \beta_2 * X_i + e_i$$

where *TN* is the predictor variable, *chlorophyll a* is the response variable, β_0 is the intercept, β_1 is the nitrogen slope, *X* is another covariate of interest, β_2 is the covariate slope, e_i is the error term, and *i* is an index for each observation or row in the data set. Multiple linear regression is simply a linear regression model with more than one predictor variable.

One statistical assumption for multiple linear regression is that the data are independent. Statistical independence means that data from different observations do not depend on each other. An example of *dependent* data is repeated measurements of blood pressure of patients over time. In this example, one would expect some level of homogeneity within each patient. This homogeneity or similarity of observations within a group (here, patient ID is the group factor) lowers the *effective* sample size of the model relative to the assumed sample size. This discrepancy can underestimate the level of uncertainty, which leads to overconfidence in the model results.

Dependent data can be accounted for in statistical models. One method is to model the group factor as a "random effect" in a linear mixed effects model. Multiple linear regression will estimate a single error variance parameter, which is an estimate of how “noisy” the model fit is relative to the observed data. In contrast, a mixed effects model will estimate an additional error variance for each specified random effect group.

In addition to accounting for data dependency, hierarchical models can use the estimated group variance to estimate the amount that each group level differs from the global estimate. For example, consider the following hierarchical model:

$$Chlorophyll_{ij} = \left(\beta_0 + \beta_{0j}\right) + \left(\beta_1 + \beta_{1j}\right) * TN_{ij} + \beta_2 * X_{ij} + e_{ij}$$

where β_0 is the global intercept, β_{0j} are the intercept adjustments for each embayment group, β_1 is the global nitrogen slope, β_{1j} are the slope adjustments for each embayment group, β_2 is the covariate slope, e_{ij} is the error term, *j* is an index for each embayment group, and *i* is an index for each observation within group *j*. Note that β_0 represents a single estimate, while β_{0j} represents *j* estimates. β_0, β_1, and β_2 are the fixed effects, and β_{0j} and β_{1j} are the random effects. The above model contains both random intercepts and random slopes.

An equivalent hierarchical representation of the above model is the following model:

$$\begin{aligned} Chlorophyll_{ij} &= \gamma_{0j} + \gamma_{1j} * TN_{ij} + \beta_2 * X + e_{ij} && \text{[Level 1 − Grab Sample Level]} \\ \gamma_{0j} &= \beta_0 + \beta_{0j} && \text{[Level 2 − Embayment Level]} \\ \gamma_{1j} &= \beta_1 + \beta_{1j} && \text{[Level 2 − Embayment Level]} \end{aligned}$$

where γ (gamma) represents the combined effects of β and β_j. This representation shows the multilevel structure of the model.

Hierarchical models are appropriate for data that can be organized at multiple levels. For this subtask, grab sample data (level 1) is available across multiple embayments (level 2). Note that open water observations were modeled separately, but the Connecticut River area of influence was included in the embayment models.

Multilevel data sets can be analyzed a number of ways. One method is to ignore the groups, pool all of the data together, and use linear regression. This pooling method provides global estimates of intercept and slope, but does not produce any insight into potential differences among the groups. Another way to analyze multilevel data is to partition the data set by group and analyze each group separately using linear regression. This partitioning method provides group-only estimates of intercept and slope. However, statistical power could be greatly reduced (i.e., uncertainty in estimates has increased) due to the smaller sample size for each group-specific model. Uncertainty could also vary greatly across models if group sample sizes are unbalanced.

Hierarchical models represent the best of both the pooled and the partitioned methods. A hierarchical model estimates both global and group-adjusted parameters in one model. One can view the group-adjusted estimates as a compromise between the global fit and the group-only fit (Figure F-2).

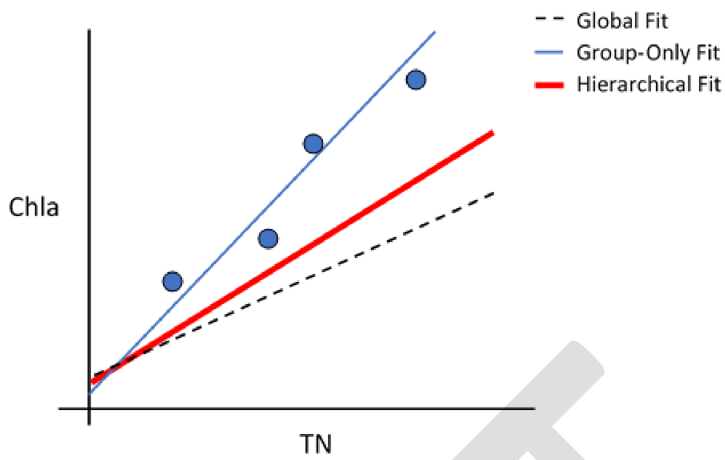

Figure F-2. Illustrative Example of How a Hierarchical Fit is a Weighted Average between the Global Fit and the Group-Only Fit

The group-adjusted or hierarchical fit estimates from a hierarchical model are known as "shrinkage" estimates. This is because the group-only estimates that are produced from partitioned models are shrunk toward the global estimate, as shown in Figure F-3. Note that Figure F-2 and Figure F-3 are complementary. In Figure F-2, one group is shown shrunk toward the global fit. In Figure F-3, all of the groups are shown shrunk toward the global estimate. Groups with fewer observations will be more heavily influenced by the global fit.

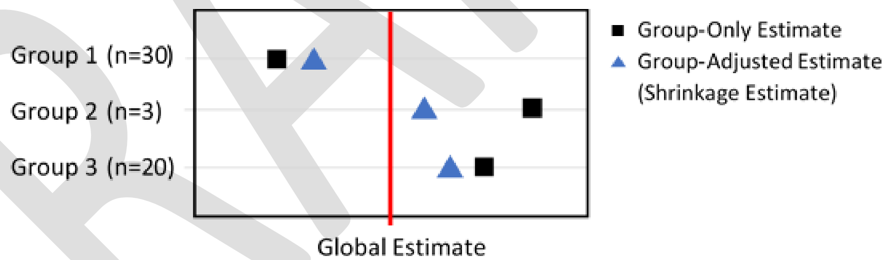

Figure F-3. Illustrative Example of how Group-Only Estimates are Shrunk toward the Global Estimate

Hierarchical models treat the groups as a random effect. Another option for analyzing multilevel data is to include the groups as a fixed effect in a linear regression model:

$$Chla_i = \beta_0 + \beta_1 * TN_i + \beta_2 * group_i + e_i$$

where β_2 is a set of group-specific intercept estimates.

Modeling the groups as fixed effects is appropriate when all of the potential group levels are known and adequately represented in the data set, and the researcher is interested in hypothesis testing for differences among the groups. For example, if the groups are "treatment" and "control" in an experiment and the goal is to see if the difference between the two groups is significant, then modeling the two groups as a fixed effect is appropriate.

However, it is often the case that all of the potential groups are not represented in the data set, or group representation is severely unbalanced in the data set, or hypothesis testing among group levels is not of interest. In an experimental study, one anticipates the treatment group to respond differently than the control group. In contrast, the underlying science for many observational studies might suggest a single global trend for the population, with some variation allowed for each group within the population. The variation among groups could be due to unobserved factors. Hierarchical models are appropriate when one assumes a global trend but also wants the flexibility to account for group variation. A hierarchical model is also better able to handle unbalanced data sets due to the model's ability to "borrow strength" from the global fit (recall Figure F-3).

LIS Embayment Hierarchical Models

Tetra Tech developed hierarchical regression models (described above) to quantify various relationships for selected embayments. Thirteen embayments and one riverine system (Connecticut River) were identified by EPA as watersheds to focus on. Data for the Connecticut River area of influence were modeled along with the embayment data, rather than modeled separately, due to the sparsity of paired data for the Connecticut River and the fact that the Connecticut River estuary is essentially an embayment. The 40-percent Connecticut River contribution isopleth (July–September) based on a particle tracking model using the NYHOPS hydrodynamic model (*Subtask E: Summary of Hydrodynamic Analysis)* was used to delineate the southern extent of the zone of influence for the mouth of the Connecticut River. Data from additional LIS embayments (beyond the embayments focused on for this analysis) were also included in the hierarchical models. The additional data allow the models to better estimate the overall LIS trend, which in turn will aid in the estimate for embayments, particularly those with few paired observations. Paired data used for the analyses is available in Appendix F4.

The goal is to develop nitrogen endpoints that protect eelgrass and other aquatic life (the assessment endpoints) within LIS. A simplified conceptual model of nutrient effects on eelgrass and other aquatic life is shown in Figure F-4. See the *Literature Review Memo* for additional information on various LIS assessment endpoints.

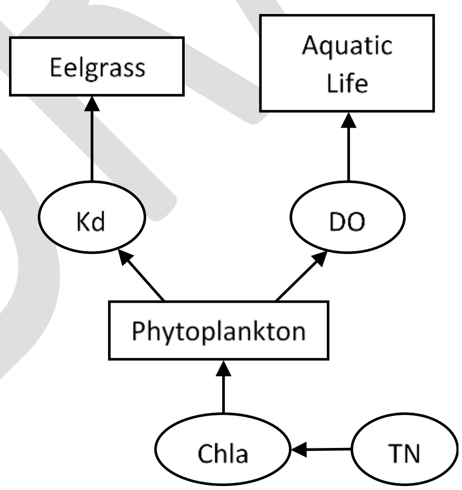

Figure F-4. Simplified Diagram Illustrating the Relationship among Key Variables and Assessment Endpoints

Eelgrass and aquatic life abundance are directly affected by light availability and DO. Eelgrasses require a sufficient quantity of surface light reaching colonization depths in order to grow. If ambient light is attenuated by particles in the water column, including algal biomass (phytoplankton), dissolved organic matter, and/or suspended inorganic sediment, then there may be insufficient light at depth for eelgrass growth and survival. The light attenuation coefficient (K_d) is a function of water clarity. Increased dissolved organic matter and particulates, including phytoplankton biomass, reduce clarity (increase light attenuation) through both absorbance and scatter of light. Phytoplankton biomass is typically quantified by measuring levels of chlorophyll, a primary algal photosynthetic pigment.

The range of percent of surface light required at maximum colonization depth for eelgrass (*Zostera marina*, the dominant seagrass in LIS) across the Northern Hemisphere ranges from 4 to 44 percent (Latimer et al. 2014), and along the East Coast of the United States, minimum requirements for eelgrass populations range between 15 and 35 percent (Latimer et al. 2014). The mean value used by Latimer et al. (2014) was 22 percent, which was also cited as a growing season average endpoint value in Vaudrey (2008). More recent long-term (more than 100 days) experimental mesocosm research found that higher light requirements were required for seedling development and growth (Ochieng et al. 2010). Seedlings grown at 34 and 58 percent surface irradiance had greater photosynthetic capacity than those grown at 11 percent. Similarly, morphological growth measures (shoot, rhizome growth, and shoot production), critical for long-term survival, were significantly higher at 34 and 58 percent than at 11 percent; however, growth at 34 percent was still less than optimal to maintain long-term meadows. The authors conclude that "seedlings exposed to light levels less than 34 percent surface irradiance during the growing season are unlikely to survive winter light and temperature stress," suggesting that light levels above 34 percent might be necessary for sufficient growth to sustain successful development of seedlings (Ochieng et al. 2010). Additional research will be needed to improve confidence in these higher endpoints, since lower percent surface irradiance endpoints have been used based on other, shorter duration and field studies for some time (Ochieng et al. 2010; Latimer et al. 2014). Balancing this information for our modeling purposes, Tetra Tech used the range of minimum requirement values from the East Coast (15 percent and 35 percent) and the mean value (22 percent) as light attenuation endpoints to protect seagrasses. Where we discuss light attenuation endpoints below, however, we discuss how higher percent light levels, including endpoints presented in Ochieng et al. (2010), could be met at various depths for different K_d values. So, while we use the endpoints above to derive K_d values, we discuss how those values might also provide higher light for seagrasses across a range of depths.

The Lambert-Beer law governs the relationship between light attenuation, depth, and percent surface light resulting in the following equation:

$$z = \frac{\ln(\frac{i_z}{i_o})}{-k_d}$$

where z is the maximum colonization depth, i_z is light at depth, i_0 is light at surface, and K_d is the light attenuation coefficient in m^{-1}.

Rearranging yields the following equation to estimate the K_d endpoint value:

$$k_d = \frac{\ln(\frac{i_z}{i_o})}{-z}$$

Tetra Tech considered values of 0.15, 0.22, and 0.35 for (i_z/i_0), percent surface light at depth. For the maximum colonization depth (z), Tetra Tech took the seagrass habitat suitability map coverages of

Vaudrey et al. (2013) and mapped them along with embayment bathymetry from Vaudrey et al. (2013). The Vaudrey bathymetry data indicated maximum depths less than 1 m and average depths less than 0.5 m for some embayments; values which may not accurately reflect ground conditions based on a review of NOAA charts. Therefore, the maximum and average depths of those embayments were set to 1 m and 0.5 m, respectively, for the purposes of Table F-2 and Table F-3. The maximum colonization depth of suitable habitat (habitat suitability scores greater than 50 based on Vaudrey et al. 2013) for each embayment was used as an estimate for the colonization depth (z). EPA explored using a minimum habitat suitability score of 88 for calculating Table F-2 and Table F-3. However, EPA is concerned with both the restoration of future seagrasses and the protection of existing seagrasses. Focusing on only areas with values above 88 would have ignored existing seagrass beds that currently maintain populations and merit protection. Only four of the embayments (Pawcatuck, Niantic, Stonington, Connecticut River) had suitable area when using a threshold value of 88; but many more embayments still maintain and support seagrasses (or are suitable to) and will continue to increase doing so as ongoing nutrient reduction efforts continue to improve habitat conditions. Moreover, the maximum colonization depths for habitats scoring > 88 were only 2.10–2.75 m, as low as 25 percent of those depths for index scores of 50, meaning that only a small portion of potential habitat would be improved. Furthermore, protecting light levels suitable to depths in areas with index values of 50 and above will inherently protect those areas with values of 88, while also potentially increasing the area with scores of 88, seeing as percent light reaching the bottom is the most weighted factor in the habitat suitability index (See Table 1 in Vaudrey et al. 2013). For these reasons, EPA used the habitat suitability target of 50 to derive light endpoints, providing greater light at the bottom to more area and increasing the potential for habitat area suitable for seagrasses. Given the values for (i_z/i_0) and (z) as described above, the values of K_d for the embayments were calculated and are presented in Table F-2. These can be compared to the values of K_d for the embayments calculated based on the average colonization depths for each embayment derived from the same bathymetry coverages used for maximum colonization depth above (Table F-3). These values show the light attenuation at average depths in the embayments.

Table F-2. Maximum Colonization Depth[a] (m) and Endpoint K_d Values for Various Light Requirements

Embayment	Maximum Colonization Depth (m)	K_d at 15% Light Requirement (m^{-1})	K_d at 22% Light Requirement (m^{-1})	K_d at 35% Light Requirement (m^{-1})
Pawcatuck River, RI and CT	-4.94	0.38	0.31	0.21
Stonington Harbor, CT	-5.19	0.37	0.29	0.20
Saugatuck River, CT	-2.00	0.95	0.76	0.52
Norwalk Harbor, CT	-2.41	0.79	0.63	0.44
Mystic River, CT	-2.61	0.73	0.58	0.40
Niantic Bay, CT	-8.75	0.22	0.17	0.12
Farm River, CT	-1.00	1.90	1.51	1.05
Southport Harbor/Sasco Brook, CT	-1.00	1.90	1.51	1.05
Northport-Centerport Harbor Complex, NY	-3.00	0.63	0.50	0.35
Port Jefferson Harbor, NY	-8.13	0.23	0.19	0.13
Nissequogue River, NY	-1.00	1.90	1.51	1.05
Stony Brook Harbor, NY	-2.00	0.95	0.76	0.52
Mt. Sinai Harbor, NY	-7.48	0.25	0.20	0.14

Embayment	Maximum Colonization Depth (m)	K_d at 15% Light Requirement (m^{-1})	K_d at 22% Light Requirement (m^{-1})	K_d at 35% Light Requirement (m^{-1})
Connecticut River, CT[b]	-8.67	0.22	0.17	0.12

[a] Bathymetry depths were for mean lower low water (low tide during the spring tidal phase).
[b] Based on suitable habitat within the 40-percent contribution isopleth identified in subtask E.

Table F-3. Average Colonization Depth[a] (m) and Endpoint K_d Values at Various Light Requirements

Embayment	Average Colonization Depth (m)	K_d at 15% Light Requirement (m^{-1})	K_d at 22% Light Requirement (m^{-1})	K_d at 35% Light Requirement (m^{-1})
Pawcatuck River, RI and CT	-0.70	2.71	2.16	1.50
Stonington Harbor, CT	-1.81	1.05	0.84	0.58
Saugatuck River, CT	-0.64	2.96	2.36	1.64
Norwalk Harbor, CT	-0.90	2.10	1.68	1.16
Mystic River, CT	-0.58	3.26	2.60	1.81
Niantic Bay, CT	-3.96	0.48	0.38	0.27
Farm River, CT	-0.50	3.79	3.03	2.10
Southport Harbor/Sasco Brook, CT	-0.50	3.79	3.03	2.10
Northport-Centerport Harbor Complex, NY	-1.08	1.76	1.40	0.97
Port Jefferson Harbor, NY	-3.25	0.58	0.47	0.32
Nissequogue River, NY	-0.50	3.79	3.03	2.10
Stony Brook Harbor, NY	-0.56	3.36	2.68	1.86
Mt. Sinai Harbor, NY	-1.62	1.17	0.93	0.65
Connecticut River, CT[b]	-1.97	0.96	0.77	0.53

[a] Bathymetry depths were for mean lower low water (low tide during the spring tidal phase).
[b] Based on suitable habitat within the 40-percent contribution isopleth identified in subtask E.

In order to model K_d as a function of chlorophyll *a*, an equation converting the value from Secchi depth (SD in meters, the most common clarity data available) to K_d was needed. There is a large range of published conversion values. Batiuk et al. (2000) provides a long list (see Table III-5 from that document) and the values for the product of K_d x SD ranged from 1.25 to 2.02 (unitless). Other values include those summarized by Koenings and Edmundson (1991) and, for clear and turbid seawater, ranged from 1.44 to 1.90. To get the conversion equation, one would divide the product by SD. Batiuk et al. (2000) recommend a value of 1.45 for Chesapeake Bay, which is consistent with turbid seawater estimates from Koenings and Edmundson (1991). Using a value of 1.45, yields the following conversion equation:

$$k_d = \frac{1.45}{SD}$$

Observed K_d values were used when available. Where SD values were available but K_d values were not, K_d values were calculated using the above equation. The maximum observed K_d value was 2.84. Calculated K_d values that were above 2.84 were removed to avoid potential extrapolation of inferred values. This approach does not account for attenuation of light by epiphyte growth on the leaf or episodic drifting algae, so it is an underestimate.

DO is, of course, required of all aerobic aquatic organisms and is now typically measured using oxygen probes. Both Connecticut and New York have existing water quality criteria for DO to protect aquatic life and these were used as initial endpoints for models of chlorophyll versus DO. These values are as follows:

- New York (class SA, SB, and SC): Chronic DO shall not be less than a daily average of 4.8 mg/L (may fall below 4.8 mg/L for a limited number of days, as defined in the formula in the water quality standards); acute DO shall not be less than 3.0 mg/L at any time.
- Connecticut (class SA and SB): Chronic DO not less than 4.8 mg/L with cumulative periods of DO in the 3.0–4.8 mg/L range as detailed in footnote in Table 1 of the water quality standards; acute DO not less than 3.0 mg/L at any time.

Because the Pawcatuck River shares a border with both Connecticut and Rhode Island, Rhode Island's DO criteria are listed below.

- Rhode Island (class SA, SA{b}, SB, SB1):
    - For surface waters above a seasonal pycnocline: instantaneous DO not less than 4.8 mg/L more than once every 3 years, except as naturally occurs.
    - For waters below the seasonal pycnocline: aquatic Life Uses are considered to be protected if conditions do not fail to meet protective thresholds, as described below, more than once every three years. Waters with a DO concentration above an instantaneous value of 4.8 mg/L shall be considered protective of aquatic life uses. When instantaneous DO values fall below 4.8 mg/L, the waters shall not be: (1) less than 2.9 mg/L for more than 24 consecutive hours during the recruitment season; nor (2) less than 1.4 mg/L for more than 1 hour more than twice during the recruitment season; nor (3) shall they exceed the cumulative DO exposure presented in Table 3.A.
    - For waters without a seasonal pycnocline, aquatic life uses are considered to be protected if conditions do not fail to meet protective thresholds, as described below, more than once every three years. DO concentrations above 4.8 mg/L shall be considered protective of aquatic life uses. When instantaneous DO values fall below 4.8 mg/L, the waters shall not be: (1) less than 3.0 mg/L for more than 24 consecutive hours during the recruitment season; nor (2) less than 1.4 mg/L for more than 1 hour more than twice during the recruitment season; nor (3) shall they exceed the cumulative DO exposure presented in Table 3.A. Cumulative low DO exposures in the 2.95–4.8 mg/L range shall be evaluated as described in Section II of the water quality standards, but shall not exceed the information presented in Table 3.B.

Table 3.A. Saltwater DO Criteria for Waters Below the Seasonal Pycnocline

24 Hour (Daily) DO Exposure Concentration (mg/L)	Daily Percent larval Mortality (%)	Allowable Number of Days Without Exceeding a 5% Reduction in Seasonal Larval Recruitment
4.6	4.96	42
4.5	6.05	30
4.4	7.36	24
4.3	8.93	20
4.2	10.79	18
4.1	12.98	16
4	15.55	14
3.9	18.51	12
3.8	21.88	10
3.7	25.69	9
3.6	29.89	8
3.5	34.47	7
3.4	39.36	6
3.3	44.46	5
3.2	49.69	4
3.1	54.92	3
3	60.05	2
2.9	64.97	1

Table 3.B. Saltwater DO Criteria for Waters without a Seasonal Pycnocline

24 Hour (Daily) DO Exposure Concentration (mg/L)	Daily Percent larval Mortality (%)	Allowable Number of Days Without Exceeding a 5% Reduction in Seasonal Larval Recruitment
4.6	4.96	16
4.5	6.05	14
4.4	7.36	12
4.3	8.93	11
4.2	10.79	10
4.1	12.98	8
4	15.55	7
3.9	18.51	6
3.8	21.88	5
3.7	25.69	4
3.6	29.89	3
3.5	34.47	2
3.4	39.36	1

Source: Rhode Department of Environmental Management. 2006. *Water Quality Regulations*. https://www.epa.gov/sites/production/files/2014-12/documents/riwqs.pdf.

Three relationships were modeled separately to determine protective nitrogen endpoints for eelgrass and other aquatic life (see Figure F-5). Note that the DO model assumes the response of bottom DO in a stratified water column. First, K_d versus chlorophyll was modeled to determine what level of chlorophyll is associated with levels of K_d that protect eelgrass habitat (left plot). Next, DO versus chlorophyll was modeled to determine what level of chlorophyll is associated with levels of DO that protect aquatic life (middle plot). The protective levels of chlorophyll identified in these first two modeling efforts were then used in the third modeling effort. Specifically, chlorophyll versus TN was modeled to determine what level of nitrogen is associated with levels of chlorophyll that protect both eelgrass and other aquatic life (right plot). The end results of the models are embayment-specific nitrogen levels that protect eelgrass and aquatic life in LIS.

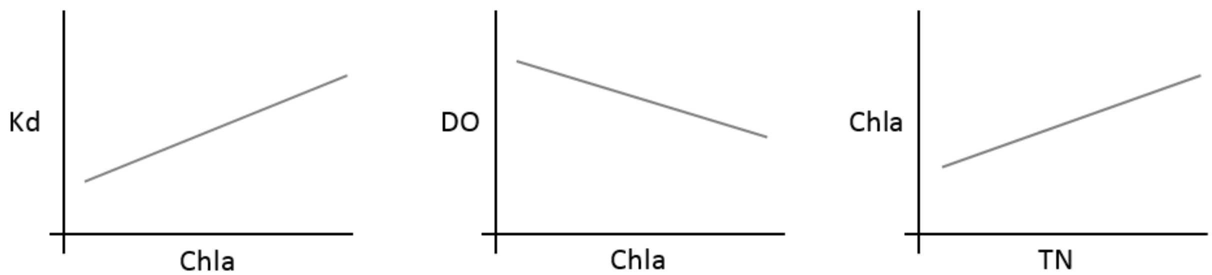

Figure F-5. Illustration of the Stressor-Response Relationships Modeled

Separate hierarchical models were developed for each of the three relationships. The data for these models are the same as described in Subtask D (*Summary of Existing Water Quality Data*), which contains observations from 14 organizations and 588 monitoring stations across 17 years. Data preparation for each model was similar. Chlorophyll *a*-corrected was used for each model rather than chlorophyll *a*, as there were more data available for the corrected measurement. Variables present in the original data set but not in the statistical model were dropped. Observations without paired data were removed (i.e., rows with missing data were removed).

Additional variables (covariates) that might help better explain the identified relationships were explored. Each additional variable added to the model has the potential to decrease the sample size due to the paired data requirement. Therefore, only variables that did not appreciably reduce the sample size were considered. Salinity, pH, and temperature were identified as covariates to consider for inclusion in the models, where applicable.

One assumption of many statistical methods is that the data are independent. Dependent data that are treated as independent will underestimate the level of uncertainty, which leads to overconfidence in the model results. To account for the repeated measurements at each station, station ID was included as a random effect—essentially a group factor—in the hierarchical models.

R software (R Core Team 2016) was used to perform the statistical analyses. The package “lme4” (Bates et al. 2015) was used to run the hierarchical models. Typical linear models assume a Gaussian (normal) distribution and an identity link function. For Subtask F, generalized linear mixed models were used. A generalized model uses additional distributions and link functions to better fit the data. Gaussian and gamma model distributions were explored. Also, identity, square root, and natural log link functions were explored. Quantile-quantile plots, residuals versus linear predictor scatterplots, observed versus fitted (1:1) plots, and residuals versus embayment boxplots were used to assess fit among candidate models. Locally weighted scatterplot smoothing (LOESS) lines were overlaid on residual plots to help identify potential trends.

After the appropriate distribution and link functions were identified, the significance of the additional covariates pH, salinity, and temperature was evaluated. Covariables with a p-value greater than 0.05 were removed, the model was rerun, and residual plots were reexamined. Natural log and square root transformations of the main predictor variable (either chlorophyll or nitrogen) were also explored.

LIGHT ATTENUATION VERSUS CHLOROPHYLL RELATIONSHIP

The relationship between K_d and chlorophyll for LIS embayment data was explored using a hierarchical model. Various model distributions and link functions were evaluated using residual plots, and covariates were evaluated using p-values. Following is the final model:

$$Kd_{ij} = (\beta_0 + \beta_{0j}) + (\beta_1 + \beta_{1j}) * \ln(Chla_{ij}) + \beta_2 * pH_i + \beta_3 * Salinity_i + \beta_4 * Temperature_i + e_{ij}$$

where β_0 is the global intercept, β_{0j} are the intercept adjustments for each embayment group, β_1 is the global chlorophyll slope, β_{1j} are the slope adjustments for each embayment group, β_2 is the pH slope, β_3 is the salinity slope, β_4 is the temperature slope, e_{ij} is the error term, *j* is an index for each embayment group, and *i* is an index for each observation within group *j*. In the above mixed effects model, “embayment groups” are random effects. Specifically, the model has a random intercept for each embayment (B_{0j}) and a random TN slope for each embayment.

Random effects for station ID were included to account for data dependency. A gamma distribution with a natural log link was used. The final model contained 1,193 observations across 78 stations. A plot of observed versus fitted values from the final model is shown in Figure F-6.

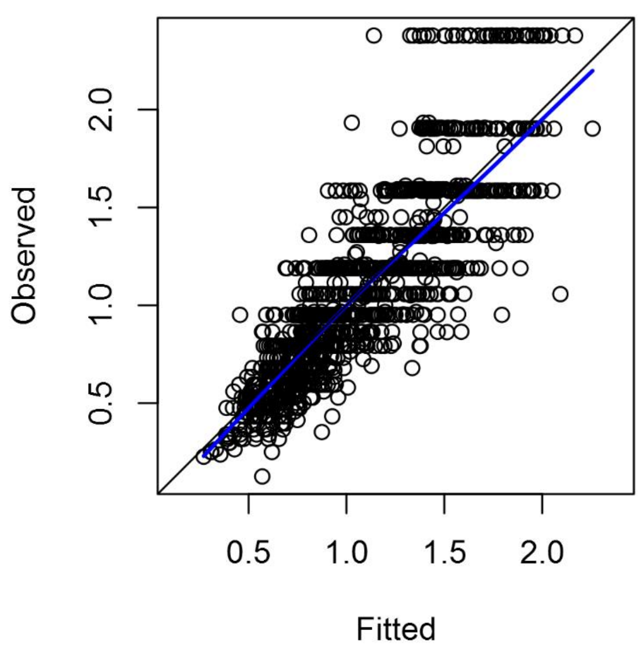

Figure F-6. Observed vs. Fitted (1:1) Plot from the K_d vs. Chlorophyll Model; the Blue Line is a LOESS Fit

The model fit the observed data reasonably well (pseudo R-squared = 0.72.), as seen in the above plot. Pseudo R-squared for the stressor-response analyses was defined as the square of r (Pearson correlation between observed and fitted values). However, K_d versus chlorophyll embayment fitted plots predicted moderate K_d levels even at extremely low chlorophyll levels. One potential explanation is that one or more factors outside of the hierarchical model is increasing K_d. It is possible that well-known contributions to K_d by suspended sediment and dissolved organic matter within the embayments were contributing to light attenuation and increasing the error in the chlorophyll versus K_d relationship. Suspended solids or dissolved organic matter data were not available and could not be modeled.

Since the individual embayment plots could not be used to derive a chlorophyll endpoint from clarity, Tetra Tech performed a linear quantile regression analysis of ln(K_d) (with K_d derived from SD using the equation above) versus ln(chlorophyll) at the 10th quantile for all the embayment data pooled together (Figure F-7). Whereas least squares and hierarchical models estimate the conditional mean of the response, quantile regression estimates the conditional quantile of the response and is advocated for use in ecological models where a response is affected by multiple factors (e.g., suspended sediment or dissolved organic matter in this case), variances are nonhomogeneous, and one seeks to understand the greatest constraint of one predictor in the absence of other factors (Cade and Noon 2003). The final quantile regression model contained 1,438 observations across 81 stations (see Table F-4). The relationship between chlorophyll *a* and K_d suggests the constraint of algal biomass on K_d is least biased (i.e., less influenced by dissolved organic matter and suspended sediment interference) at lower quantiles (Figure F-7). This bias is likely lower at the lower quantiles because chlorophyll places the greatest constraints on K_d when minimally influenced by these other factors. After looking at the 5th and 20th quantiles, the 10th quantile was selected as sufficiently characteristic of the unbiased relationship while also containing enough values to be reasonably estimated. K_d endpoint values of 0.5 and 0.7 were selected for interpolation. A K_d endpoint value of 0.5 equates to 15 percent light at 4 m, 22 percent light at 3 m, 35 percent light at 2m, and 50 percent light at 1.4 m. This value provides from 14 to 78 percent surface light (mean: 57 percent) across the range of average embayment depths (min: 0.5 m, max: 4.0 m) and 52 percent at the mean average embayment depth (1.3 m) (Table F-3). It provides an average of 26 percent light across the range of maximum colonization depths (Table F-2). The value of 0.7 is the K_d endpoint recommended for LIS (Yarish et al. 2006; Vaudrey 2008), a value corresponding to 40 percent of water column surface light available to 1.3 m, the average of primary tier average colonization depths. This value provides from 6 to 70 percent surface light (mean: 48 percent) across the range of average embayment colonization depths (Table F-3). The chlorophyll values associated with these K_d endpoints using the 10th quantile regression model are listed in

Table F-5.

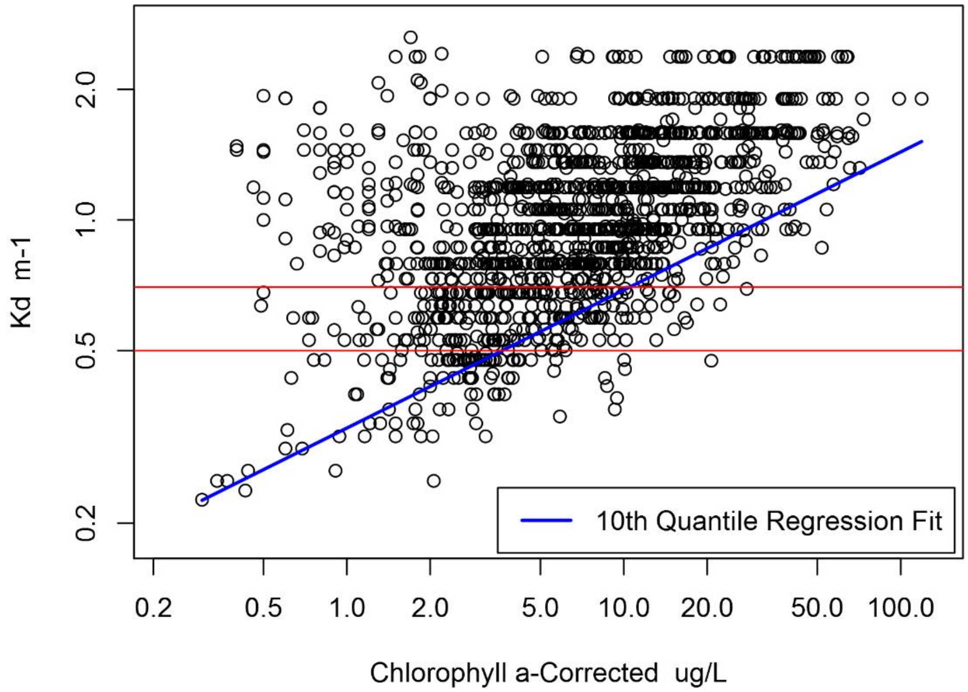

Figure F-7. K_d vs. Chlorophyll 10th Quantile Regression Relationship for LIS (Plotted in Natural Log Transformed Space)

Table F-4. Paired Observations and Station Counts for the K_d vs Chlorophyll Embayment Model, by Embayment

Embayment Name	Observation Count	Station Count
Pawcatuck River, RI and CT	234	6
Stonington Harbor, CT	-	-
Saugatuck Estuary, CT	4	2
Norwalk Harbor, CT	-	-
Mystic Harbor, CT	-	-
Niantic Bay, CT	2	1
Farm River, CT	-	-
Southport Harbor / Sasco Brook, CT	-	-
Northport-Centerport Harbor Complex, NY	176	9
Port Jefferson Harbor, NY	256	13
Nissequogue River, NY	23	3
Stony Brook Harbor, NY	64	7
Mt. Sinai Harbor, NY	39	4
Connecticut River, CT	-	-
Other Waters	640	36
Total	**1,438**	**81**

Table F-5. K_d Endpoints and Associated Chlorophyll *a* Values, based on 10th Quantile Regression Model

K_d Endpoint	Associated Chlorophyll *a* Value Based on 10th Quantile
0.5	3.71
0.7	9.83

DO VERSUS CHLOROPHYLL RELATIONSHIP

The relationship between DO and chlorophyll for LIS embayment data was explored using a hierarchical model. Various model distributions and link functions were evaluated using residual plots, and covariates were evaluated using p-values. Following is the final model:

$$DO_{ij} = \left(\beta_0 + \beta_{0j}\right) + \left(\beta_1 + \beta_{1j}\right) * sqrt(Chla_{ij}) + \beta_2 * Temperature_i + e_{ij}$$

where β_0 is the global intercept, β_{0j} are the intercept adjustments for each embayment group, β_1 is the global chlorophyll slope, β_{1j} are the slope adjustments for each embayment group, β_2 is the temperature slope, e_{ij} is the error term, *j* is an index for each embayment group, and *i* is an index for each observation within group *j*. In the above mixed effects model, “embayment groups” are random effects. Specifically, the model has a random intercept for each embayment (B_{0j}) and a random TN slope for each embayment.

Random effects for station ID were included to account for data dependency. A Gaussian distribution with an identity link was used. The final model contained 1,828 observations across 124 stations (see Table F-6). A plot of observed versus fitted values from the final model is shown in Figure F-8.

Table F-6. Paired Observations and Station Counts for the Dissolved Oxygen vs Chlorophyll Embayment Model, by Embayment

Embayment Name	Observation Count	Station Count
Pawcatuck River, RI and CT	362	11
Stonington Harbor, CT	49	5
Saugatuck Estuary, CT	8	4
Norwalk Harbor, CT	-	-
Mystic Harbor, CT	77	2
Niantic Bay, CT	4	2
Farm River, CT	-	-
Southport Harbor / Sasco Brook, CT	-	-
Northport-Centerport Harbor Complex, NY	176	9
Port Jefferson Harbor, NY	245	13
Nissequogue River, NY	27	6
Stony Brook Harbor, NY	64	7
Mt. Sinai Harbor, NY	42	6
Connecticut River, CT	-	-
Other Waters	774	59
Total	**1,828**	**124**

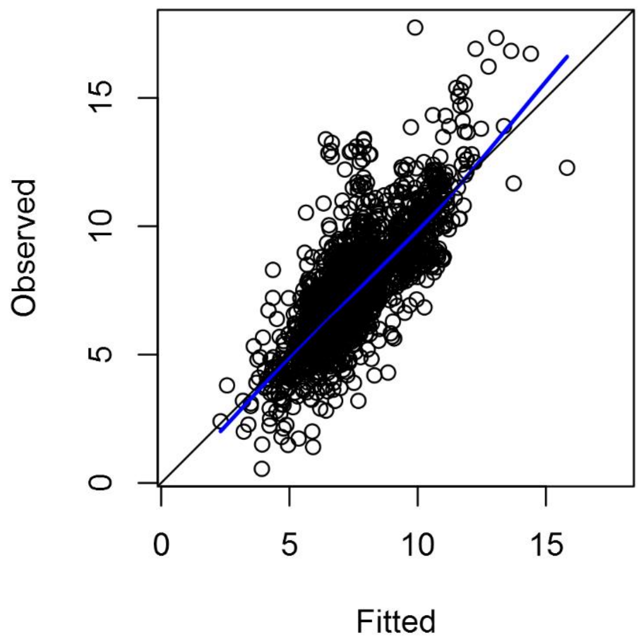

Figure F-8. Observed vs. Fitted (1:1) Plot from the DO vs Chlorophyll Model (Blue Line is a LOESS Fit)

The model fit reasonably well (pseudo R-squared = 0.61), as seen in the above plot. However, the coefficient for chlorophyll was positive, suggesting that DO levels increase as chlorophyll increases. Also, the fitted model predicted relatively high values of DO even at extremely low chlorophyll levels.

The relationship between DO and phytoplankton is complex. Phytoplankton can contribute to DO via photosynthesis and can deplete DO through respiration. DO levels, as a result, fluctuate daily, tending to be highest in late afternoon and lowest in early morning. Grab samples of DO are, therefore, of little utility in gaging the complete manifestation of metabolic effects on DO. Profile DO levels can be more informative during stratification, but need to be paired with chlorophyll data. There were sparse data available for paired samples taken at the bottom of the water column across LIS (40 observations). These daytime grab sample DO values, therefore, likely obscure the effects of metabolism on DO. In addition, other factors also affect DO levels. For example, it is possible that the DO levels of the embayments are largely influenced by the embayment's residence time, mixing with the open sound, and reaeration. For these reasons, a chlorophyll endpoint was not able to be derived for the DO versus chlorophyll relationship.

CHLOROPHYLL VERSUS NITROGEN RELATIONSHIP

The relationship between chlorophyll and TN for LIS embayment data was explored using a hierarchical model. Various model distributions and link functions were evaluated using residual plots, and covariates were evaluated using p-values. Following is the final model:

$$Chlorophyll_{ij} = \left(\beta_0 + \beta_{0j}\right) + \left(\beta_1 + \beta_{1j}\right) * TN_{ij} + \beta_2 * pH_i + \beta_3 * Temperature_i + e_{ij}$$

where β_0 is the global intercept, β_{0j} are the intercept adjustments for each embayment group, β_1 is the global nitrogen slope, β_{1j} are the slope adjustments for each embayment group, β_2 is the pH slope, β_3 is the temperature slope, e_{ij} is the error term, j is an index for each embayment group, and i is an index for each observation within group *j*. In the above mixed effects model, “embayment groups” are random effects. Specifically, the model has a random intercept for each embayment (B_{0j}) and a random TN slope for each embayment.

Random effects for station ID were included to account for data dependency. A gamma distribution with a natural log link was used. The final model contained 1,335 observations across 113 stations (see Table F-7). A plot of observed versus fitted values is shown in Figure F-9 (pseudo R-squared = 0.47). The model had trouble fitting three observations with high levels of chlorophyll (left plot) but performed well on the bulk of the data (right plot), although some underestimation of large chlorophyll values is still present.

Table F-7. Paired Observations and Station Counts for the Chlorophyll vs Total Nitrogen Embayment Model, by Embayment

Embayment Name	Observation Counts	Station Counts
Pawcatuck River, RI and CT	137	11
Stonington Harbor, CT	23	5
Saugatuck Estuary, CT	6	4
Norwalk Harbor, CT	-	-
Mystic Harbor, CT	35	2
Niantic Bay, CT	3	2
Farm River, CT	-	-
Southport Harbor / Sasco Brook, CT	-	-
Northport-Centerport Harbor Complex, NY	171	9
Port Jefferson Harbor, NY	221	12
Nissequogue River, NY	26	5
Stony Brook Harbor, NY	64	7
Mt. Sinai Harbor, NY	40	6
Connecticut River, CT	-	-
Other Waters	609	50
Total	**1,335**	**113**

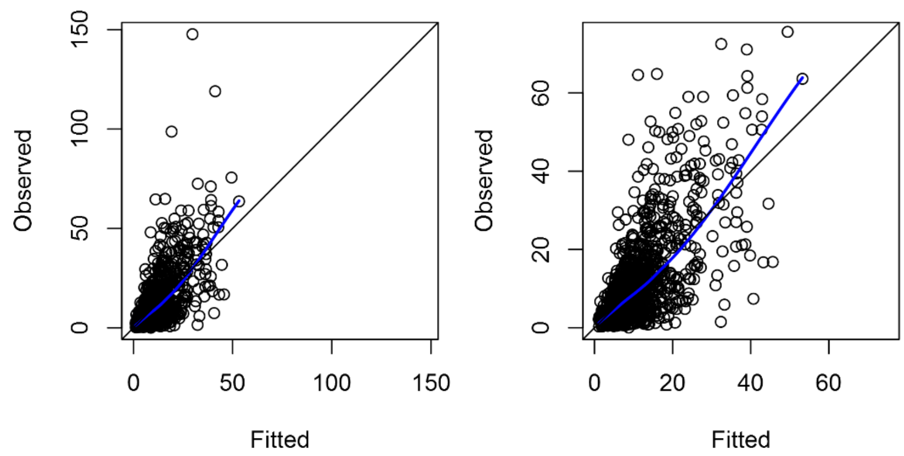

Figure F-9. Observed vs. Fitted (1:1) Plot from the Chlorophyll vs. Total Nitrogen Model (Full Plot Left); Magnified Plot (Right); Blue Line is a LOESS Fit

Tetra Tech constructed embayment-specific plots and solved for TN concentrations associated with three chlorophyll endpoints: 3.5 µg/L., 5.5 µg/L, and 10 µg/L. The first concentration was derived from the chlorophyll versus light attenuation models described above using a value of K_d = 0.5, and the second is the recommended chlorophyll endpoint for LIS provided by Vaudrey (2008). The last chlorophyll endpoint was derived from the chlorophyll versus light attenuation model described above using a value of K_d = 0.7 (9.83 rounded to 10 µg/L), the recommended K_d endpoint for LIS provided by Vaudrey (2008). Fitted plots and 90-percent confidence intervals of the chlorophyll versus nitrogen relationship are presented in *Subtask G: Nitrogen Endpoints* for each embayment. Confidence intervals in a statistical model represent the range of uncertainty in the average chlorophyll values (y-axis) the model would predict for the given TN value (x-axis). The uncertainty in the predicted values stems from the uncertainty in the estimated model parameters. The confidence intervals will vary per embayment due to the use of hierarchical modeling and the varying covariate (pH and temperature) values for each embayment.

LIS Open Water Models

In addition to the embayment models, Tetra Tech developed regression models to quantify various relationships for the Western LIS open water portion of LIS. The Western LIS is defined as the Western Narrows and the Eastern Narrows portion of LIS (recall Figure F-1). Paired data for Eastern LIS was limited (n=31; 2.5% of the sample) and was not included in this analysis. Paired data for bottom samples was also limited (n=1; <1%), so only surface observations were included in this analysis.

One extreme data point of chlorophyll *a*-corrected (106.8 µg/L) was identified. This value is over two times larger than the second largest value (46.1 µg/L). The open water models were run with and without the extreme observation, and differences in results were negligible. The extreme observation was removed from the final open water datasets.

The methodology used for the open water models followed the one outlined for the embayment models, with some exceptions. The assessment endpoint eelgrass was not considered in the deeper open water region. Also, there were only two potential group levels—Western Narrows and Eastern Narrows—for the open water data. In contrast, there were approximately 28 different embayments for the embayment models. For hierarchical models, we required that the group factor have at least five levels. Fewer levels would make it challenging for the model to accurately estimate the group variance. Therefore, there was no grouping variable in the model.

R software was used to perform the statistical analyses. The package "lme4" was also used to run the open water models. While a hierarchical analysis was not required to model shrinkage estimates, the open water models still have the same data dependency issues that are present in the embayment models. To account for the repeated measurements at each station, station ID was included as a random effect in the open water models.

DO VERSUS CHLOROPHYLL RELATIONSHIP

The relationship between DO and chlorophyll for LIS open water data was explored using the following model. Various model distributions and link functions were evaluated using residual plots, and covariates were evaluated using p-values. Following is the final model:

$$DO_i = \beta_0 + \beta_1 * Chlorophyll_i + \beta_2 * pH_i \\ + \beta_3 * Salinity + \beta_4 * Temperature_i + e_i$$

where β_0 is the intercept, β_1 is the chlorophyll slope, β_2 is the pH slope, β_3 is the salinity slope, β_4 is the temperature slope, e_i is the error term, and i is an index for each observation.

Random effects for station ID were included to account for data dependency. A gamma distribution with an identity link was used. The final model contained 1,320 observations across 25 stations (see Table F-8). A plot of observed versus fitted values is shown in Figure F-10.

Table F-8. Paired Observations and Station Counts for the Dissolved Oxygen vs Chlorophyll Open Water Model, by Open Water Group

Open Water Group	Observation Counts	Station Counts
Open Water - Western Narrows	1,198	13
Open Water - Eastern Narrows	122	12
Total	**1,320**	**25**

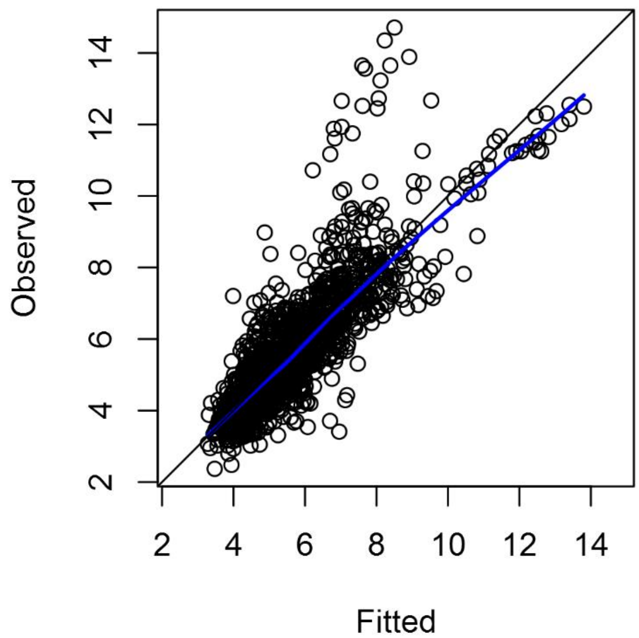

Figure F-10. Observed vs. Fitted (1:1) Plot from the DO vs. Chlorophyll Model (Blue Line is a LOESS Fit)

The model fit reasonably well (pseudo R-squared = 0.70), as seen in the above plot. However, the open water DO versus chlorophyll model results were similar to the embayment model in that the coefficient for chlorophyll was positive, suggesting that DO levels increase as chlorophyll increases. The fitted model predicted relatively high values of DO even at extremely low chlorophyll levels. The lack of paired bottom DO samples with chlorophyll data was a limitation. There was plenty of bottom DO data, as evidenced by the hypoxia maps drawn for LIS, but this analysis was unable to find adequate paired bottom DO with chlorophyll samples to build this relationship. Therefore, a chlorophyll endpoint was not able to be derived for the DO versus chlorophyll relationship.

CHLOROPHYLL VERSUS NITROGEN RELATIONSHIP

The relationship between chlorophyll and nitrogen for LIS open water data was explored using the following model. Various model distributions and link functions were evaluated using residual plots, and covariates were evaluated using p-values. Following is the final model:

$$Chlorophyll_i = \beta_0 + \beta_1 * TN_i + \beta_2 * pH + \beta_3 * Salinity + e_i$$

where β_0 is the intercept, β_1 is the chlorophyll slope, β_2 is the pH slope, β_3 is the salinity slope, e_i is the error term, and *i* is an index for each observation.

Random effects for station ID were included to account for data dependency. A gamma distribution with an identity link was used. The final model contained 1,157 observations across 18 stations (see Table F-9). A plot of observed versus fitted values is shown in Figure F-11.

Table F-9. Paired Observations and Station Counts for the Chlorophyll vs Total Nitrogen Open Water Model, by Open Water Group

Open Water Group	Observation Counts	Station Counts
Open Water - Western Narrows	1,108	11
Open Water - Eastern Narrows	49	7
Total	**1,157**	**18**

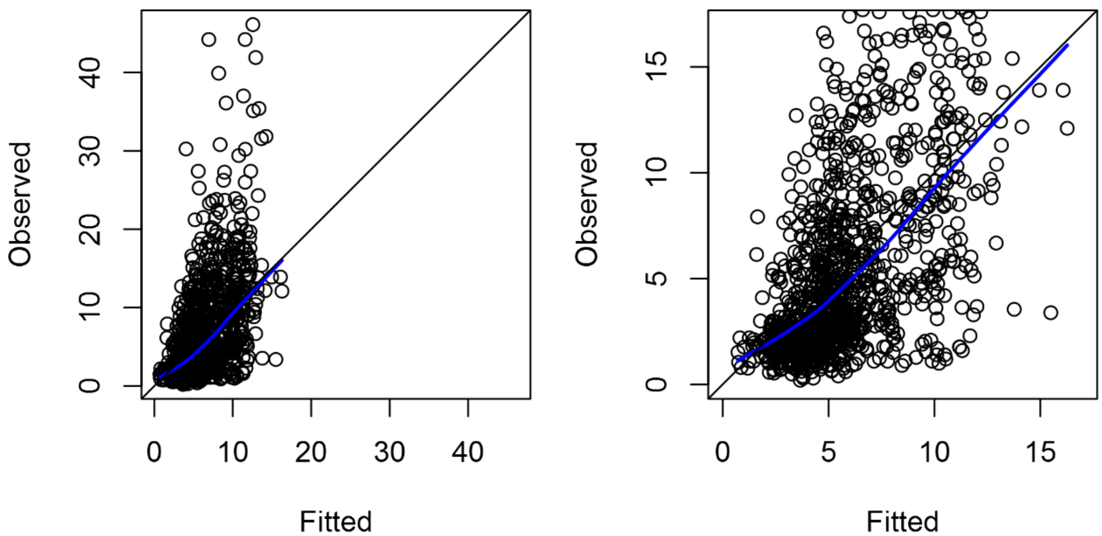

Figure F-11. Observed vs. Fitted (1:1) Plot from the Chlorophyll vs. Nitrogen Model (Full Plot Left); Magnified Plot (Right); (Blue Line is a LOESS Fit)

The results of this model were subpar (pseudo R-squared = 0.32), based on the above 1:1 plot. Additional variables were explored to see if the model fit could be improved. Temperature was found to be insignificant and was removed from the model. In addition, the coefficient for nitrogen was negative, suggesting that chlorophyll levels decrease as nitrogen increases. Therefore, a nitrogen endpoint was not able to be derived for the chlorophyll versus nitrogen relationship for the open waters.

Distribution-Based Analysis

There are multiple methods for using a distribution-based analysis to develop nutrient endpoints. The results of distribution-based analysis are being applied in this application within a multiple lines of evidence approach that includes scientific literature and empirical stressor-response models linking nitrogen (concentrations and/or loads) to response variables in the different watersheds.

The distribution-based line of evidence refers to evaluating distributions of nitrogen concentrations in different watersheds and using those values to inform protective nitrogen endpoints. EPA has used the distribution of nutrient concentrations from minimally disturbed reference watersheds for setting nutrient endpoints for a number of applications, including TMDLs and permitting (USEPA 1999, 2001, 2015, 2016). The same concept can be extended to distributions from other nutrient concentration populations as well, including those from time periods known to be supporting uses (i.e., temporal reference) (USEPA 2010) and from populations known to be supporting their designated uses, especially aquatic life uses (USEPA 2015). In this way, identifying distributions of nitrogen concentrations or loads for embayments known to exhibit good water quality conditions can provide a line of evidence for those values that protect valued uses and, thus, provide a line of evidence for developing nitrogen endpoints.

The term *reference condition* refers to the condition that supports biological integrity, defined as “the capability of supporting and maintaining a balanced, integrated, adaptive community of organisms having a species composition, diversity, and functional organization comparable to that of the natural habitat of the region” (Frey 1977). Another definition of *integrity* is defined as “the ability of an aquatic community to support and maintain a structural and functional performance comparable to the natural habitats of a region” (Frey 1977; Karr and Dudley 1981). Because the ideal reference condition for biological integrity rarely exists due to the history of human disturbance and existing pervasive impacts (e.g., atmospheric pollutants), biological integrity *sensu strictu* must often be estimated and the ideal reference condition replaced with a surrogate.

Distribution-based approaches have traditionally been used to estimate reference conditions for nutrients in two ways: (1) as an upper percentile of least disturbed reference conditions or conditions known to be supporting uses and (2) a lower percentile of distributions representing all similar waters. In the former case, an upper percentile is used because the distribution is expected to include waters meeting uses and the corresponding nutrient conditions are expected to be supportive. EPA’s recommendation in this case has been the 75th percentile (USEPA 2001). In the latter case, the lower percentile is used because the distribution is expected to contain degraded waters given that it is an entire population. This latter option is most useful in regions where the number of legitimate natural reference water bodies is usually very small, such as highly developed land-use areas like those surrounding LIS. EPA’s recommendation in this case is usually the lower 25th percentile (USEPA 2001). In either case, the selection of percentile should reflect confidence in the degree of degradation represented by either population. If almost all waters are impacted to some extent, then the lower 5th percentile might be used to approximate previous natural conditions. The actual distribution of the observations and knowledge of the inherent regional water quality are also determinants of the percentile selected.

Other elements that must be considered in interpreting distributions are where and when the data were gathered. If the sample size is large enough, the time of year the individual samples were taken may not matter; either all seasons will be represented or most of the data will cluster about an appropriate index season. Similarly, surface grab or depth-selected samples or composite samples might not matter if the diverse data set is large enough. For this application, the same depth criteria used for the stressor-response analysis were applied here; for timeframe, the entire range of dates was used given the size of the available data set, but focus was placed on the growing season for consistency with other lines of evidence.

Methods

Tetra Tech used the 25th percentile of all samples in the approach described above based on existing guidance (USEPA 2001). EPA guidance suggests the lower quartile (25th percentile) of all water quality data be used where nutrient enrichment is present and the 75th percentile of reference waters be used where replicate near-pristine conditions are present (USEPA 2001). Given the long history of enrichment in LIS, the listing and ongoing implementation of a TMDL for nitrogen, and present knowledge related to continuing nutrient impacts in LIS that are the basis for this current work, use of the 75th percentile of reference waters seemed indefensible given that such waters would be difficult if not impossible to accurately identify or represent.

Water quality data (*Subtask D: Summary of Existing Water Quality Data*) were extracted, geometric mean seasonal (April–September) surface water TN averages for each year at a station calculated using natural log transformed TN values, and cumulative distribution functions of TN station average concentrations calculated along with distribution statistics, including the 25th percentile values for embayment waters and open waters. These values were then used as the distribution-based endpoints for embayments and open water areas.

As additional support, Tetra Tech identified two embayments (Niantic Bay and Mystic Harbor) within which seagrass coverage increased consistently between 2002 and 2012 (Tiner et al. 2013). Nutrient data from within these watersheds were compiled and reviewed for concentrations as supporting information.

Results

Stations for the distribution-based analysis were spread across embayments and open water (Figure F-12). The data from these stations were extracted and the distribution-based values estimated as described above. Distributions of values in embayments were higher than those for open water, as expected (Figure F-13). The distributional statistics are given in Table F-10.

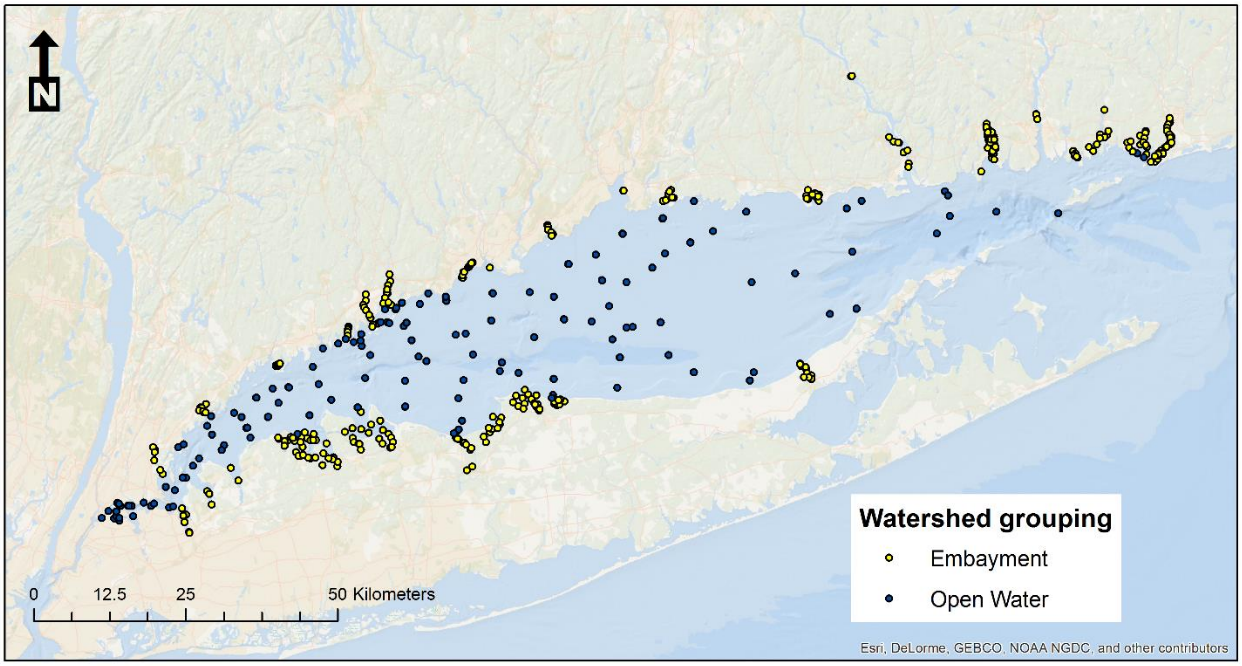

Figure F-12. Water Quality Stations Used in Distribution-Based Analysis

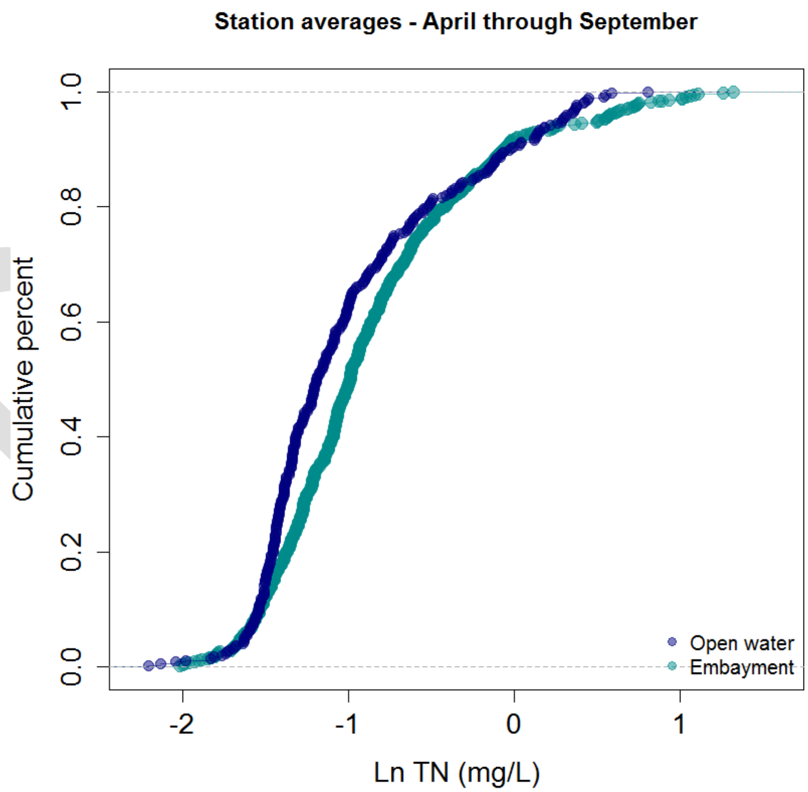

Figure F-13. Cumulative Distribution of Station Year Seasonal (April-September) ln(Total Nitrogen, mg/L) Values in Open Water versus Embayment Stations

Table F-10. Distributional Statistics of Total Nitrogen (mg/L) for the Different Distribution-Based Populations (Values in Bold are Based on Previous EPA Applications and Represent the 25th Percentile of all Sample Population Station_Year Seasonal [April-September] Log-Normal Averages)

Watershed Grouping								
Percentile	5%	10%	25%	Median	75%	90%	95%	N
All Embayments	0.19	0.22	**0.27**	0.37	0.56	0.95	1.66	587
All Open Water	0.20	0.21	**0.24**	0.30	0.50	0.98	1.34	345

As additional supporting information for estimating TN concentrations consistent with those known to be supporting desired conditions in LIS embayments, the median TN concentrations in Niantic Bay and Mystic River embayments, both of which were found to have exhibited areal seagrass increases from 2002 to 2012 (Tiner et al. 2013), were 0.21 mg/L and 0.53 mg/L (N=6 and N=61), respectively, based on available surface water quality data. The lower of these values was consistent with the 25th percentile from the distribution-based values above (0.262 mg/L), especially for Niantic Bay. This supports the conclusion that the distribution based values here are consistent with those known to support increases in seagrass.

Subtask F Sources Cited

Bates, D.M. Maechler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1):1-48. doi:10.18637/jss.v067.i01.

Batiuk, R., P. Bergstrom, M. Kemp, and M. Teichberg. 2000. *Chesapeake Bay Submerged Aquatic Vegetation Water Quality and Habitat-Based Requirements and Restoration Targets: A Second Technical Synthesis*. Printed by the United States Environmental Protection Agency for the Chesapeake Bay Program, Annapolis, MD.

Benson, J.L., D. Schlezinger, and B.L. Howes. 2013. Relationship between nitrogen concentration, light, and *Zostera marina* habitat quality and survival in southeastern Massachusetts estuaries. *Journal of Environmental Management* 131:129–137.

Cade, B.S., and B.R. Noon. 2003. A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment* 1:412–420.

Frey, D.G. 1977. The Integrity of Water: An Historical Approach. In *Proceedings of the Integrity of Water Symposium*, March 10-12, 1975, Washington, DC.

Harding, L.W. Jr., R.A. Batiuk, T.R. Fisher, C.L. Gallegos, T.C. Malone, W.D. Miller, M.R. Mulholland, H.W. Paerl, E.S. Perry, and P. Tango. 2014. Scientific bases for numerical chlorophyll criteria in Chesapeake Bay. *Estuaries and Coasts* 37:134–148.

Howes, B.L., R. Samimy, and B. Dudley. 2003. *Site-Specific Nitrogen Thresholds for Southeastern Massachusetts Embayments: Critical Indicators Interim Report*. Prepared by Massachusetts Estuaries Project for the Massachusetts Department of Environmental Protection. Accessed February 2017. http://yosemite.epa.gov/OA/EAB_WEB_Docket.nsf/Verity%20View/DE93FF445FFADF1285257527005AD4A9/$File/Memorandum%20in%20Opposition%20...89.pdf.

Karr, J.R., and D.R. Dudley. 1981. Ecological perspective on water quality goals. *Environmental Management* 5(1):55-68.

Koenings, J.P., and J.A. Edmundson. 1991. Secchi disk and photometer estimates of light regimes in Alaskan lakes: Effects of yellow color and turbidity. *Limnology and Oceanography* 36:91-105.

Latimer, J.S., M.A. Tedesco, R.L. Swanson, C. Yarish, P.E. Stacey, and C. Garza, ed. 2014. *Long Island Sound: Prospects for the Urban Sea*. Springer Series on Environmental Management, Springer-Verlag, New York.

Ochieng, C.A., F.T. Short, and D.I. Walker. 2010. Photosynthetic and morphological responses of eelgrass (*Zostera marina L.*) to a gradient of light conditions. *Journal of Experimental Marine Biology and Ecology* 382(2):117–124.

MEP. 2017. *The Massachusetts Estuaries Project: Reports Available to Download*. Downloadable individual reports for the 33 embayment systems. Massachusetts Estuary Program. Accessed February 2017. http://www.oceanscience.net/estuaries/reports.htm.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Tiner, R., K. McGuckin, and A. MacLachlan. 2013. *2012 Eelgrass Survey for Eastern Long Island Sound, Connecticut and New York: National Wetlands Inventory Report*. U.S. Fish and Wildlife Service, National Wetlands Inventory Program, Northeast Region, Hadley, MA.

USEPA. 1999. *Protocol for Developing Nutrient TMDLs*. EPA 841-B-99-007. U.S. Environmental Protection Agency, Washington DC.

USEPA. 2001. *Nutrient Criteria Technical Guidance Manual: Estuarine and Coastal Marine Waters*. EPA-822-B-01-003. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA. 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). U.S. Environmental Protection Agency, Washington DC. Accessed January 2017. https://yosemite.epa.gov/sab/sabproduct.nsf/c91996cd39a82f648525742400690127/c439b7c63eb911f8525773b004e53ca!OpenDocument.

USEPA. 2015. *Authorization to Discharge under the National Pollutant Discharge Elimination System: The City of Taunton, Massachusetts, Department of Public Works*. NPDES Permit # MA0100897. U.S. Environmental Protection Agency. Accessed February 2017. https://www3.epa.gov/region1/npdes/permits/2015/finalma0100897permit.pdf.

USEPA. 2016. *Revised Fact Sheet: EPA Proposes to Reissue a National Pollutant Discharge Elimination System (NPDES) Permit to Discharge Pollutants Pursuant to the Provisions of the Clean Water Act (CWA), City of Sandpoint Wastewater Treatment Plant, NPDES Permit #ID0020842*. U.S. Environmental Protection Agency. Accessed January 2017. https://www3.epa.gov/region10/pdf/permits/npdes/id/sandpoint_revisedFS_ID0020842_04192016.pdf.

Vaudrey, J.M.P. 2008. *Establishing Restoration Objectives for Eelgrass in Long Island Sound. Part I: Review of the Seagrass Literature Relevant to Long Island Sound.* Final Grant Report to the Connecticut

Department of Environmental Protection, Bureau of Water Protection and Land Reuse and the U.S. Environmental Protection Agency.

Vaudrey, J.M.P., J. Eddings, C. Pickerell, L. Brousseau., and C. Yarish. 2013. *Development and Application of a GIS-based Long Island Sound Eelgrass Habitat Suitability Index Model*. Final report submitted to the New England Interstate Water Pollution Control Commission and the Long Island Sound Study. 171 p. + appendices.

Yarish, C., R. E. Linden, G. Capriulo, E. W. Koch, S. Beer, J. Rehnberg, R. Troy, E. A. Morales, F. R. Trainor, M. DiGiacomo-Cohen, and R. Lewis. 2006. *Environmental Monitoring, Seagrass Mapping and Biotechnology as Means of Fisheries Habitat Enhancement along the Connecticut Coast*. Final Grant Report to CT DEP Long Island Sound Research Fund, CWF-314-R.

Subtask G. Nitrogen Endpoints

Introduction

To calculate TN endpoints, Tetra Tech used multiple lines of evidence from scientific literature, stressor-response, and distribution-based values to the maximum extent provided by the data. Tetra Tech considered each individual line of evidence equally and developed an endpoint using each one. These endpoints consisted of (1) the calculated median TN value of the literature-based values protective of seagrass endpoints, (2) the 25th percentile distribution-based TN value, and (3) the mean TN value associated with chlorophyll endpoints using the stressor-response models.

Uncertainty around the literature review line of evidence endpoints was estimated using the minimum and maximum values from Table F-1. For the embayments, Tetra Tech used literature review values for seagrass protection (range of 0.30–0.50 mg/L; median of 0.39 mg/L rounded to 0.40 mg/L); for open water, Tetra Tech used literature values for all endpoints (range of 0.30–0.60 mg/L; median of 0.41 mg/L rounded to 0.40 mg/L).

For the distribution-based line of evidence, Tetra Tech selected the 25th percentile of all samples, as further described in Subtask F and presented in Table F-10. The TN endpoints using this approach were 0.27 mg/L (all embayments) and 0.24 mg/L (open water). No uncertainty estimates around these values were calculated.

The stressor-response line of evidence was developed based on relationships between TN and chlorophyll. As a reminder, chlorophyll *a*-corrected endpoints for embayments were developed from stressor-response models of chlorophyll and light levels (K_d) necessary to protect seagrasses in embayments. A literature based chlorophyll a endpoint was also used. Stressor-response models of chlorophyll and DO were not significant for embayments or open water, and therefore no chlorophyll endpoints were generated from that analysis. The stressor-response uncertainty ranges in the tables for each embayment were the 90th percentile confidence intervals around the modeled endpoint value limited to the highest (2.52 mg/L) and lowest (0.06 mg/L) observed TN values in the empirical LIS data set.

G.1 Pawcatuck River, RI and CT

Figure G-1 shows a map of the Pawcatuck River watershed. Paired data for the embayment included 137 observations across 11 water quality stations within the growing season (April–September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll a-corrected, TN, pH, and temperature where available.

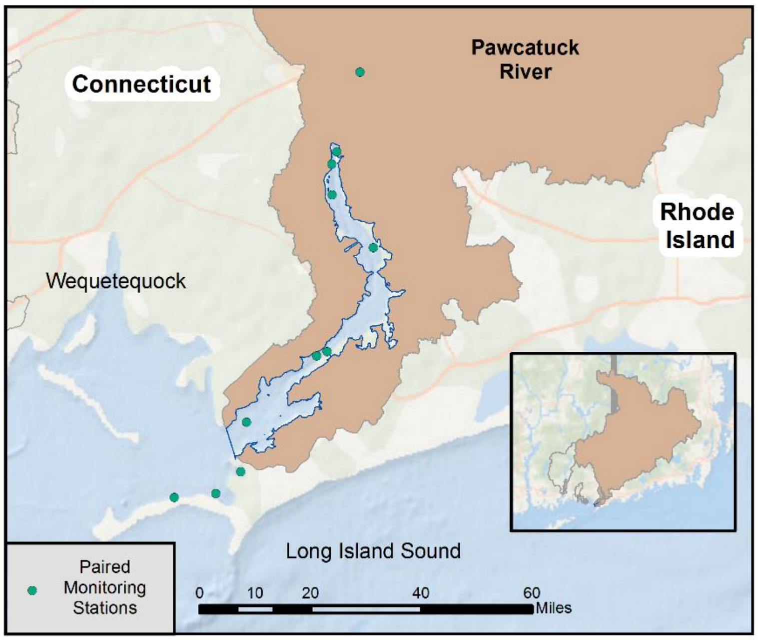

Figure G-1. Pawcatuck River Watershed, RI and CT

Chlorophyll endpoints and TN endpoints for the Pawcatuck River watershed are presented in Table G-1. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-2.

Table G-1. Chlorophyll and TN Endpoints for Pawcatuck River, RI and CT

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model Mean for Individual Embayment (90th Percent Confidence Interval)	10	**0.53**[a] (0.30–0.74)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

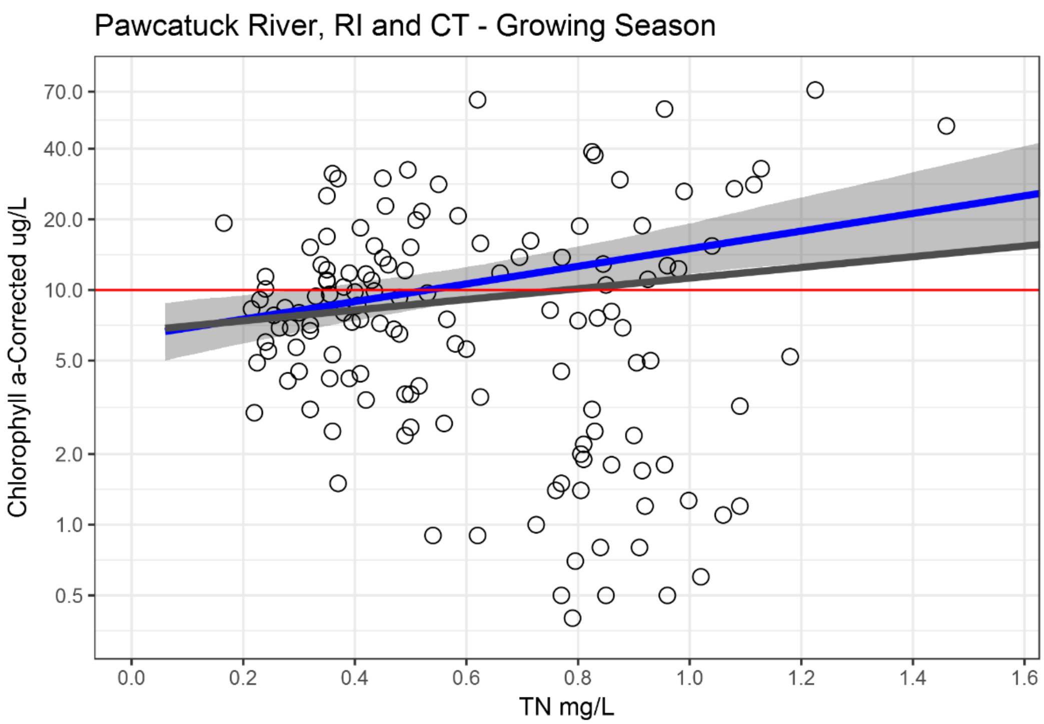

Figure G-2. Chlorophyll vs. Total Nitrogen Relationship for Pawcatuck River Watershed, RI and CT (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are provided in Table G-1.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 0.53 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 10 µg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.2 Stonington Harbor, CT

Figure G-3 shows a map of the Stonington Harbor watershed. Paired data for the embayment included 23 observations across five water quality stations within the growing season (April–September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

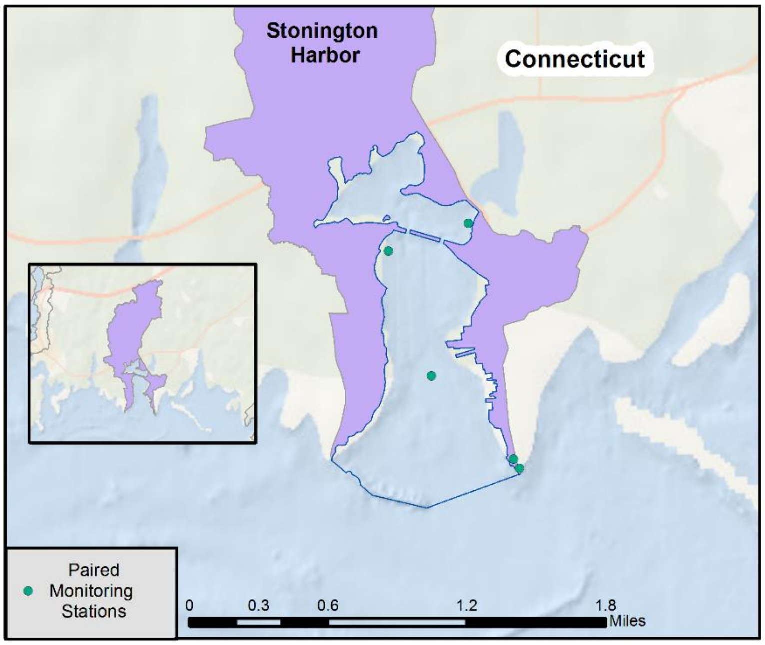

Figure G-3. Stonington Harbor Watershed, CT

Chlorophyll endpoints and TN endpoints for the Stonington Harbor watershed are presented in Table G-2. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-4.

Table G-2. Chlorophyll and TN Endpoints for Stonington Harbor Watershed, CT

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	5.5	**0.48** (0.06–2.52)
		10	**1.25**[a] (0.71–2.52)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

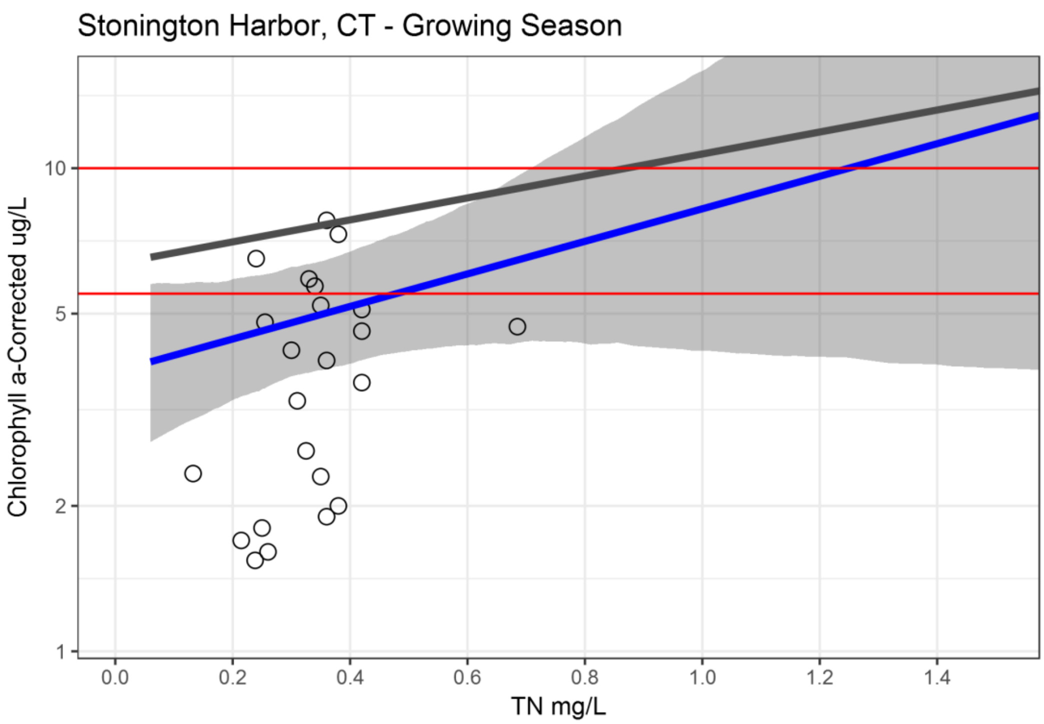

Figure G-4. Chlorophyll vs. Total Nitrogen Relationship for the Stonington Harbor Watershed, CT (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-2.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoints to protect light levels needed to support seagrasses are 0.48 mg/L and 1.25 mg/L. Respectively, these values are based on the embayment-specific chlorophyll *a* endpoints of 5.5 μg/L (consistent with the value Vaudrey (2008) recommended for LIS) and 10 μg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). These values are further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.3 Saugatuck Estuary, CT[2]

Figure G-5 shows a map of the Saugatuck Estuary watershed. Paired data for the embayment included six observations across four water quality stations within the growing season (April–September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

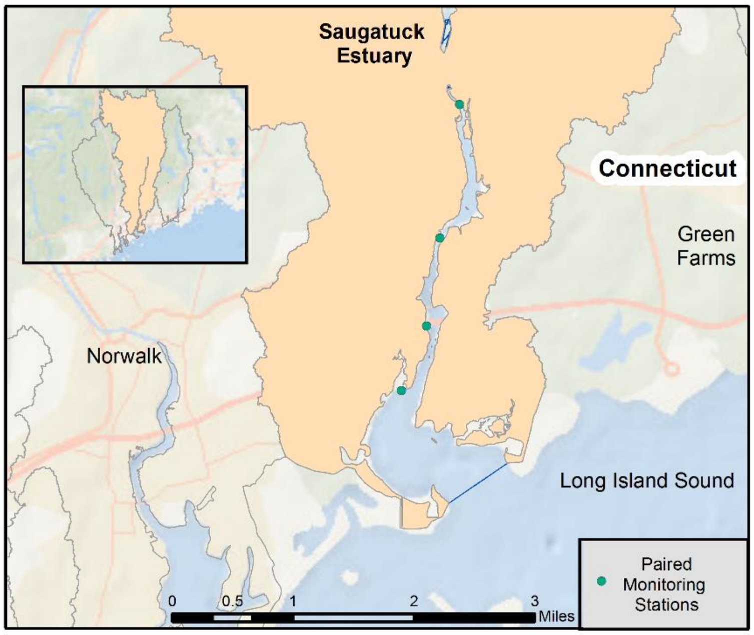

Figure G-5. Saugatuck Estuary Watershed, CT

Chlorophyll endpoints and TN endpoints for the Saugatuck Estuary watershed are presented in Table G-3. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-6.

Table G-3. Chlorophyll and TN Endpoints for Saugatuck Estuary Watershed, CT

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	10	**0.52**[a] (0.06–2.52)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

[2] Includes two Vaudrey et al. (2016) embayments: Saugatuck River, CT, and Saugatuck River, North, CT (freshwater).

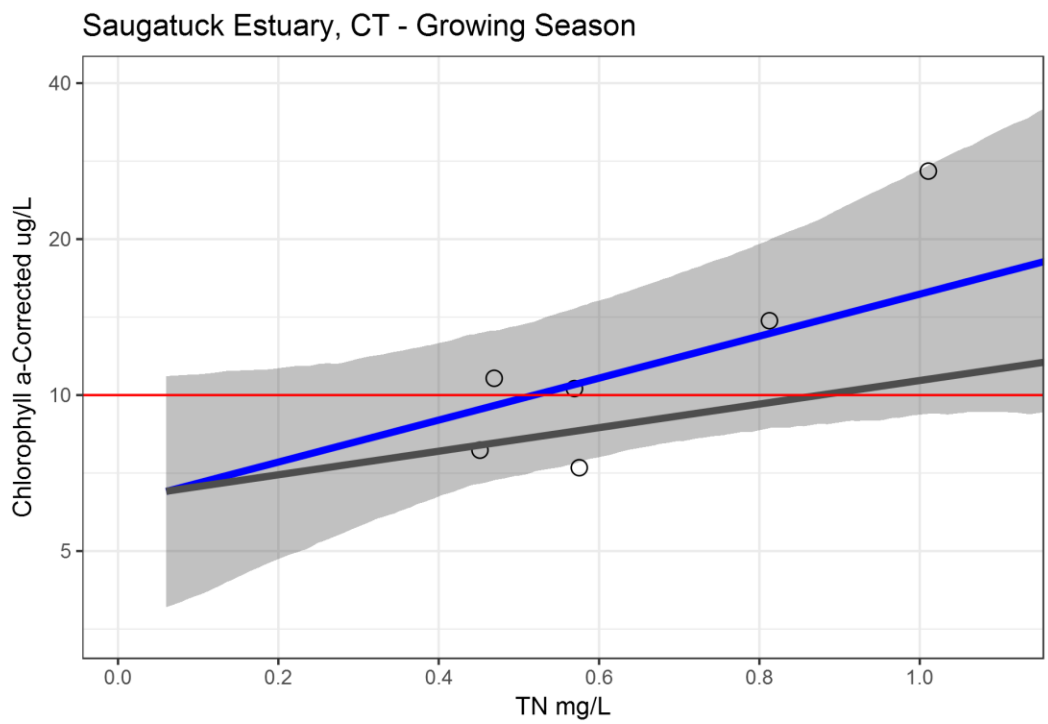

Figure G-6. Chlorophyll vs. Total Nitrogen Relationship for the Saugatuck Estuary Watershed, CT (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-3.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 0.52 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 10 µg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.4 Norwalk Harbor, CT

Figure G-7 shows a map of the Norwalk Harbor watershed. No paired data were available for the embayment within the growing season (April–September). Therefore, the global fit using data from all 1,335 embayment observations LIS-wide was used for the stressor-response analysis. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

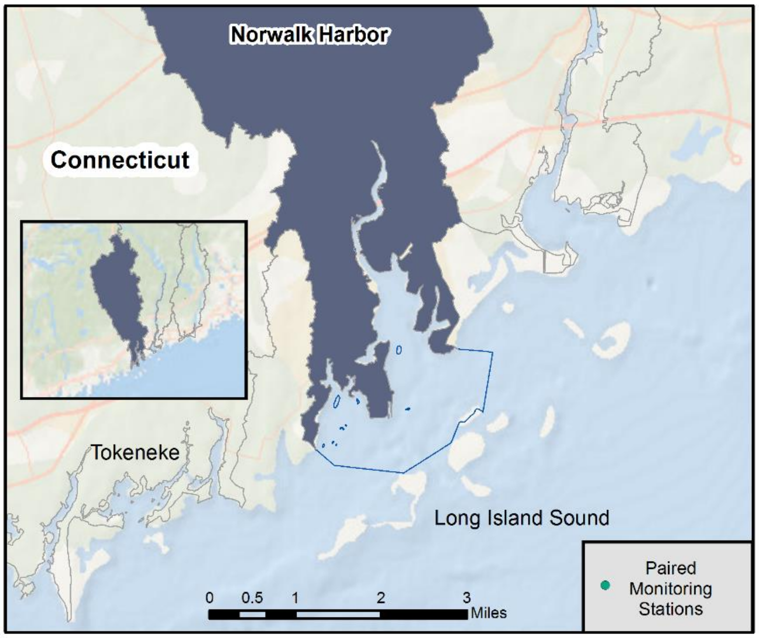

Figure G-7. Norwalk Harbor Watershed, CT

Chlorophyll endpoints and TN endpoints for the Norwalk Harbor watershed are presented in Table G-4. A plot of the chlorophyll versus TN relationship is presented in Figure G-8.

Table G-4. Chlorophyll and TN Endpoints for Norwalk Harbor Watershed, CT

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	10	**0.61**[a] (0.50–0.72)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

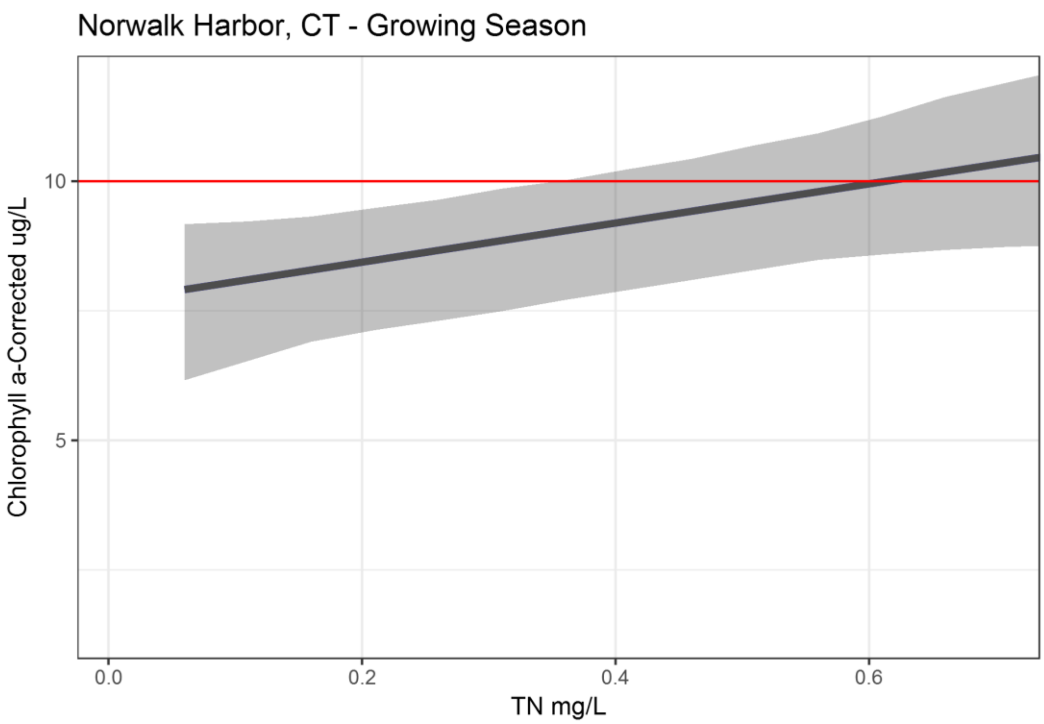

Figure G-8. Chlorophyll vs. Total Nitrogen Relationship for the Norwalk Harbor Watershed, CT (There Were No Paired Growing Season Observations Available for this Embayment; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-4.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 0.61 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 10 µg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.5 Mystic Harbor, CT

Figure G-9 shows a map of the Mystic Harbor watershed. Paired data for the embayment included 35 observations across two water quality stations within the growing season (April–September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

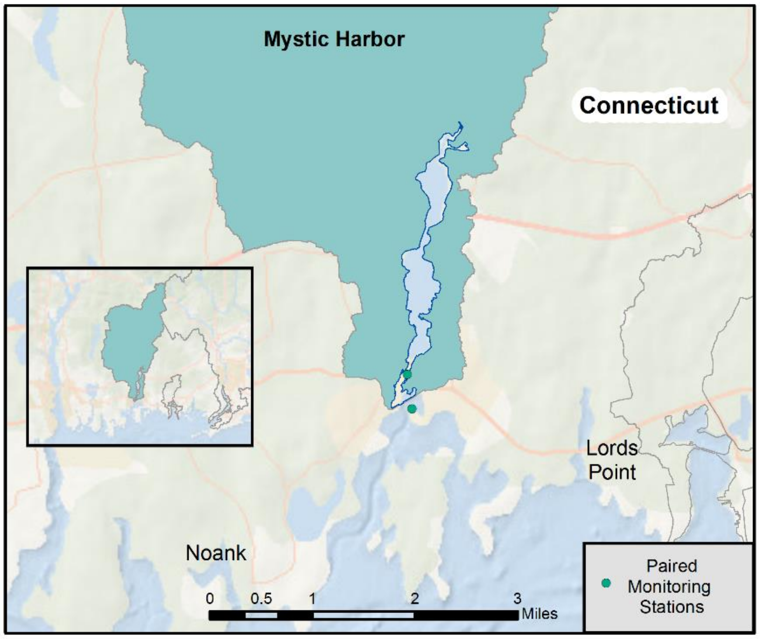

Figure G-9. Mystic Harbor Watershed, CT

Chlorophyll endpoints and TN endpoints for the Mystic Harbor watershed are presented in Table G-5. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-10.

Table G-5. Chlorophyll and TN Endpoints for Mystic Harbor Watershed, CT

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	10	**0.57**[a] (0.26–1.02)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

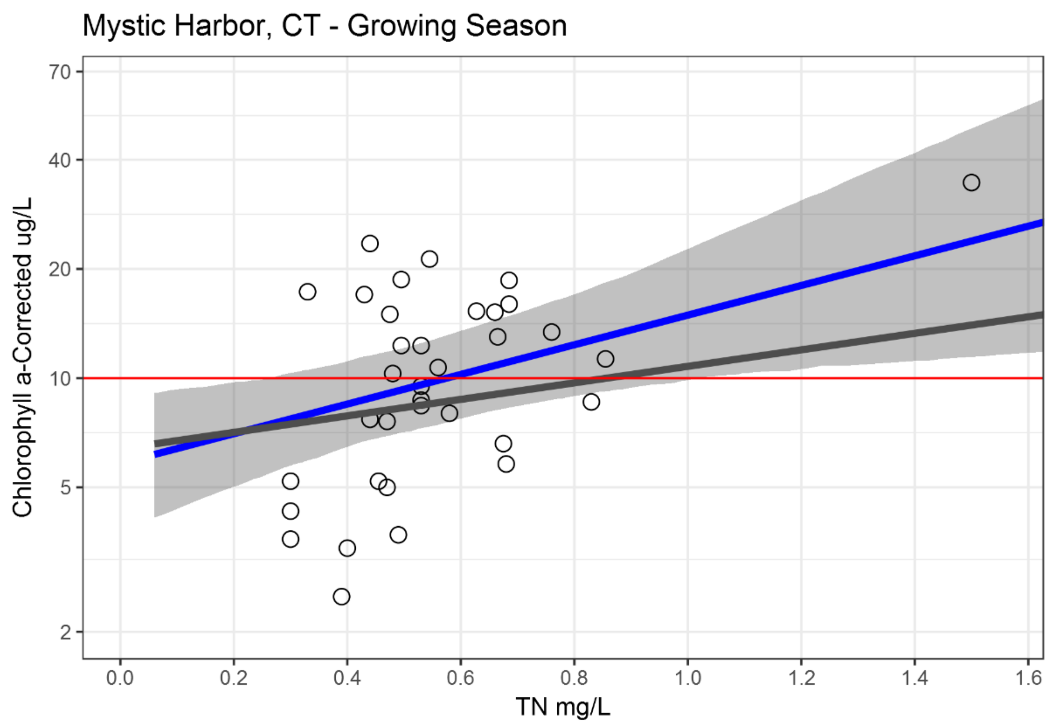

Figure G-10. Chlorophyll vs. Total Nitrogen Relationship for the Mystic Harbor Watershed, CT (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-5.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 0.57 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 10 μg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.6 Niantic Bay, CT[3]

Figure G-11 shows a map of the Niantic Bay watershed. Paired data for the embayment included three observations across two water quality stations within the growing season (April–September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

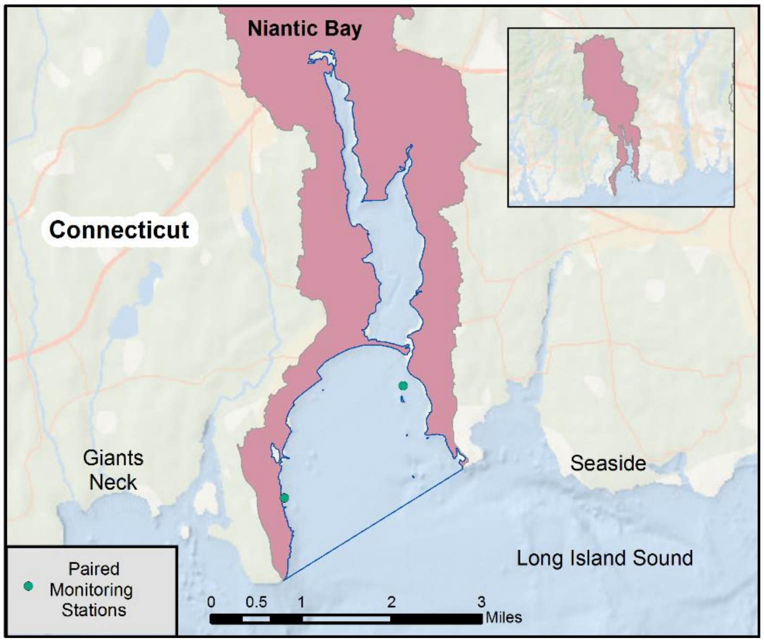

Figure G-11. Niantic Bay Watershed, CT

Chlorophyll endpoints and TN endpoints for the Niantic Bay watershed are presented in Table G-6. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-12.

Table G-6. Chlorophyll and TN Endpoints for the Niantic Bay Watershed, CT

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	10	**0.68**[a] (0.06–2.52)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

[3] Includes two Vaudrey et al. (2016) embayments: Niantic River, CT, and Niantic Bay, CT.

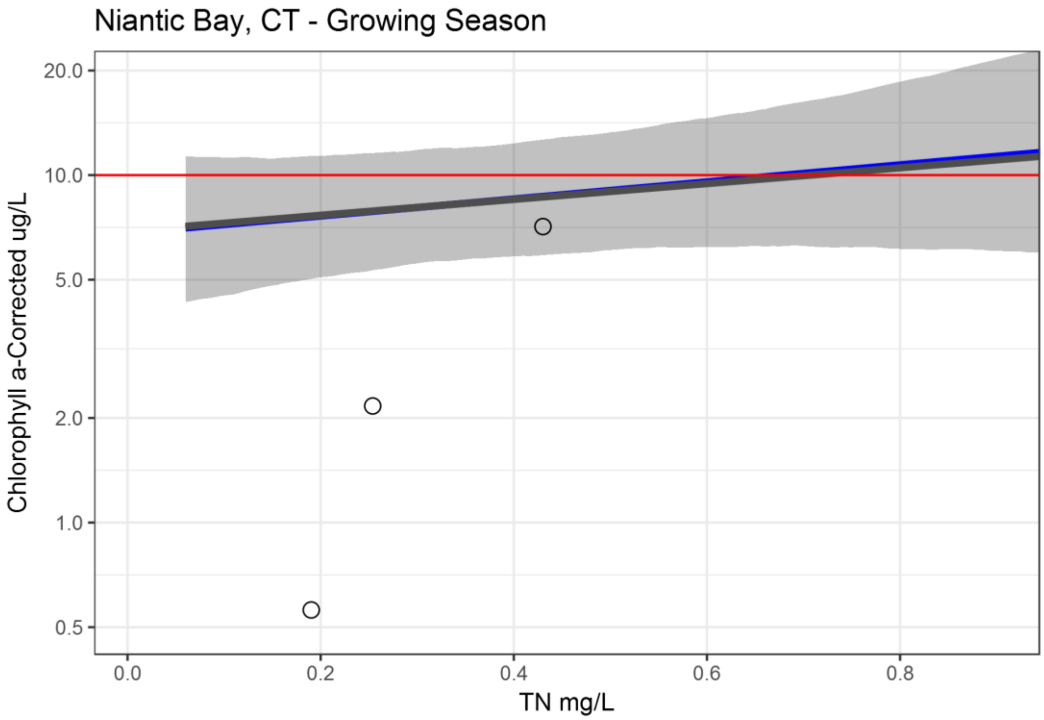

Figure G-12. Chlorophyll vs. Total Nitrogen Relationship for the Niantic Bay Watershed, CT (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-6.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 0.68 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 10 µg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.7 Farm River, CT

Figure G-13 shows a map of the Farm River watershed. No paired data was available for the embayment within the growing season (April–September). Therefore, the global fit using data from all 1,335 embayment observations LIS-wide was used for the stressor-response analysis. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

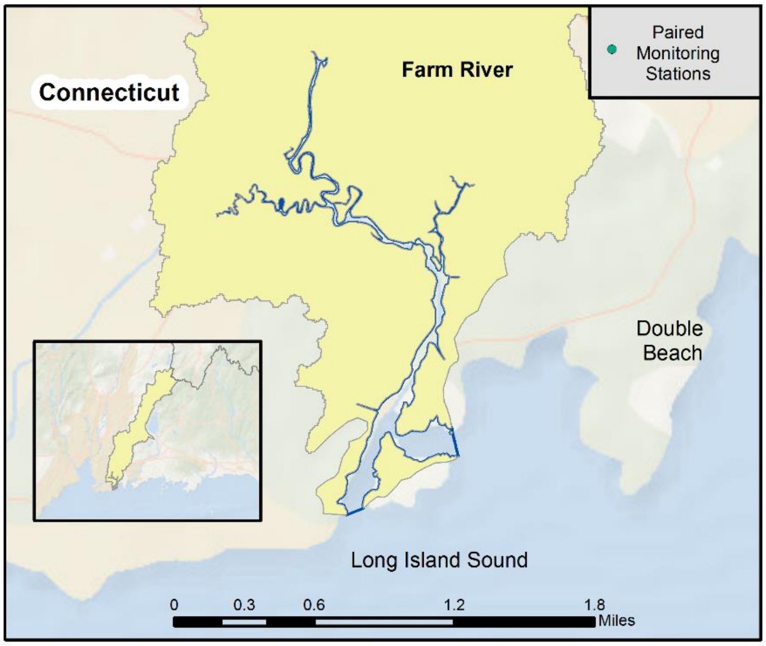

Figure G-13. Farm River Watershed, CT

Chlorophyll endpoints and TN endpoints for the Farm River watershed are presented in Table G-7. A plot of the chlorophyll versus TN relationship is presented in Figure G-14.

Table G-7. Chlorophyll and TN Endpoints for the Farm River Watershed, CT

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	10	**0.87**[a] (0.76–0.97)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

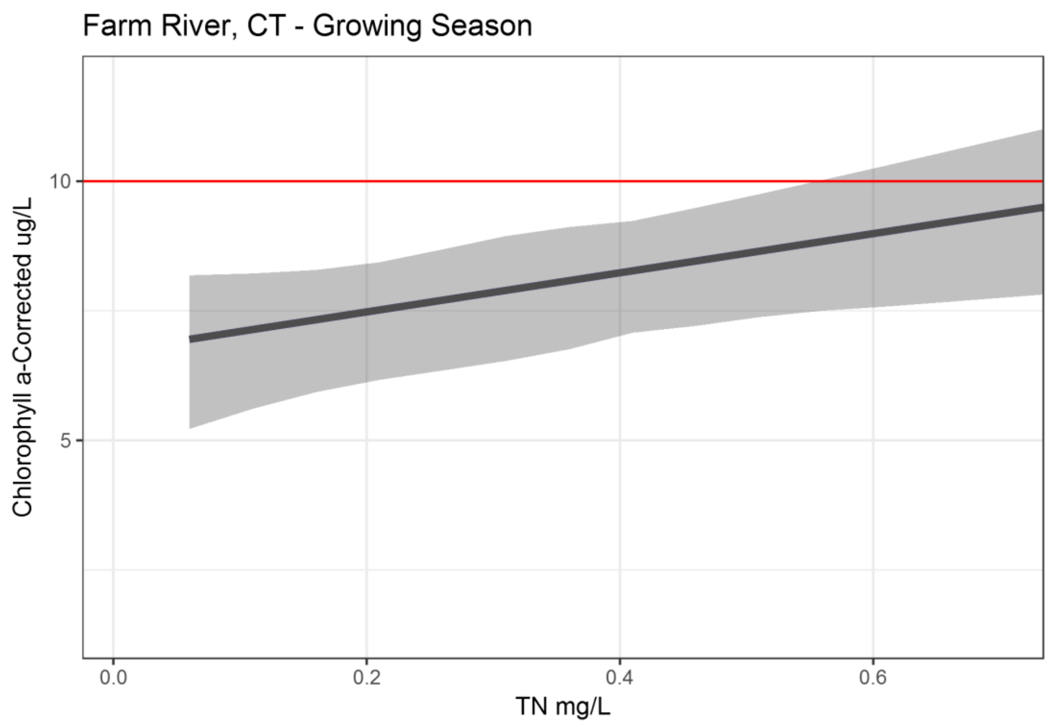

Figure G-14. Chlorophyll vs. Total Nitrogen Relationship for the Farm River Watershed, CT (There Were No Paired Growing Season Observations Available for this Embayment; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-7.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 0.87 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 10 µg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.8 Southport Harbor/ Sasco Brook, CT[4]

Figure G-15 shows a map of the Southport Harbor/ Sasco Brook watershed. No paired data was available for the embayment within the growing season (April–September). Therefore, the global fit using data from all 1,335 embayment observations LIS-wide was used for the stressor-response analysis. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

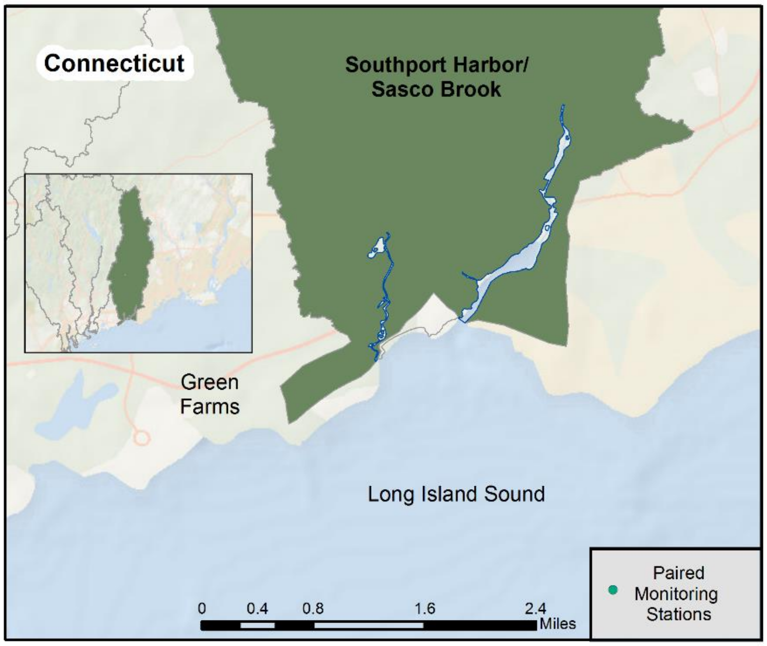

Figure G-15. Southport Harbor/Sasco Brook Watershed, CT

Chlorophyll endpoints and TN endpoints for the Southport Harbor/Sasco Brook watershed are presented in Table G-8. A plot of the chlorophyll versus TN relationship is presented in Figure G-16.

Table G-8. Chlorophyll and TN Endpoints for the Southport Harbor/Sasco Brook Watershed, CT

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	10	**0.87**[a] (0.76–0.97)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

[4] Includes two Vaudrey et al. (2016) embayments: Mill River, CT, and Sasco Brook, CT.

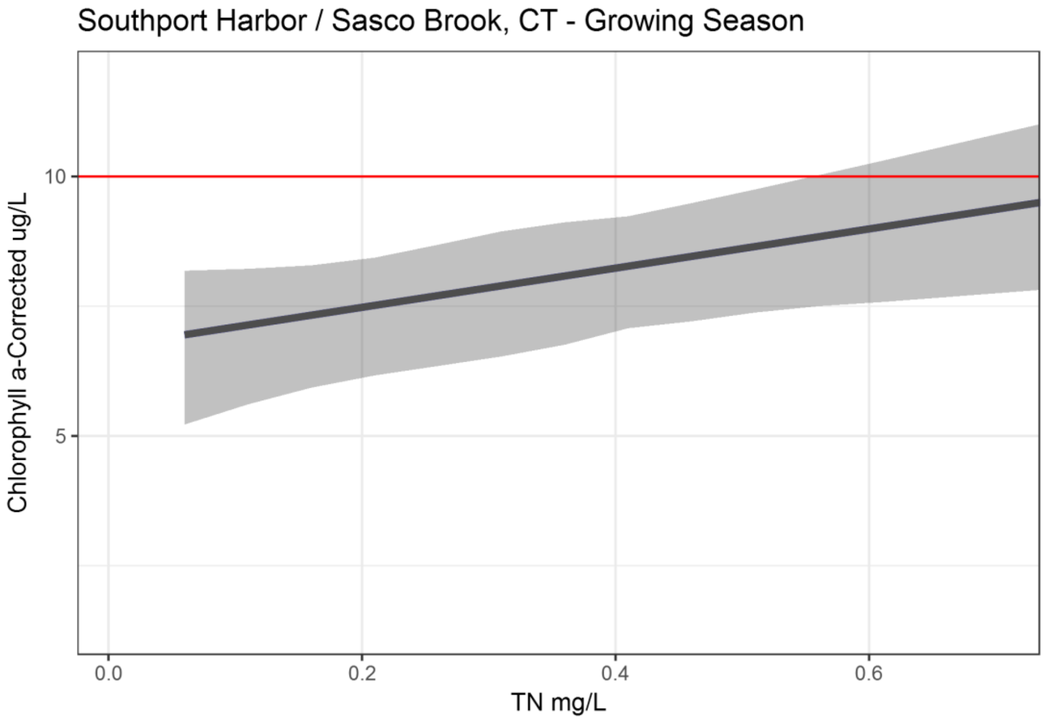

Figure G-16. Chlorophyll vs. Total Nitrogen Relationship for the Southport Harbor/Sasco Brook Watershed, CT (There Were No Paired Growing Season Observations Available for this Embayment; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-8.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 0.87 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 10 μg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.9 Northport-Centerport Harbor Complex, NY[5]

Figure G-17 shows a map of the Northport-Centerport Harbor Complex watershed. Paired data for the embayment included 171 observations across nine water quality stations within the growing season (April–September). These data were obtained from the water quality data used to for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

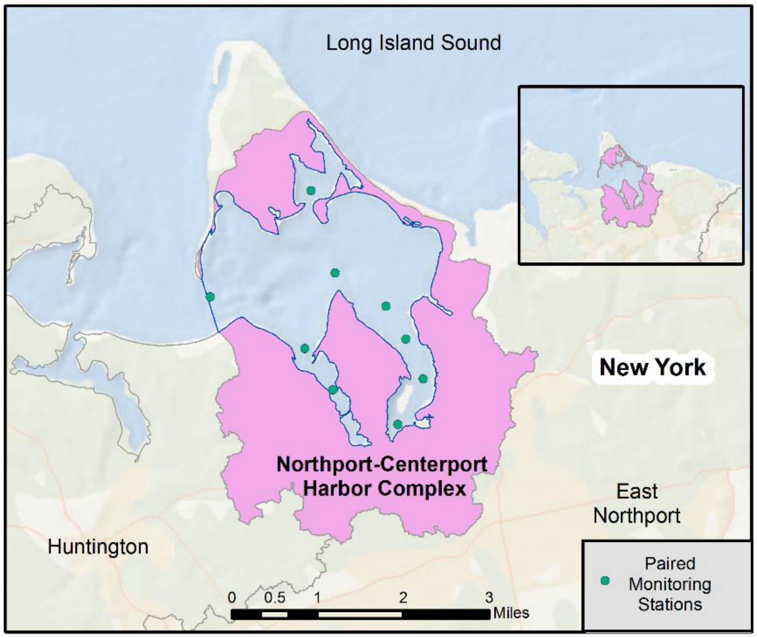

Figure G-17. Northport-Centerport Harbor Complex Watershed, NY

Chlorophyll endpoints and TN endpoints for the Northport-Centerport Harbor Complex watershed are presented in Table G-9. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-18.

Table G-9. Chlorophyll and TN Endpoints for the Northport-Centerport Harbor Complex Watershed, NY

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	5.5	**0.07[a]** (0.06–1.17)
		10	**1.27[a]** (0.74–2.52)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[5] Includes three Vaudrey et al. (2016) embayments: Centerport Harbor, NY; Northport Bay, NY; and Northport Harbor, NY.

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

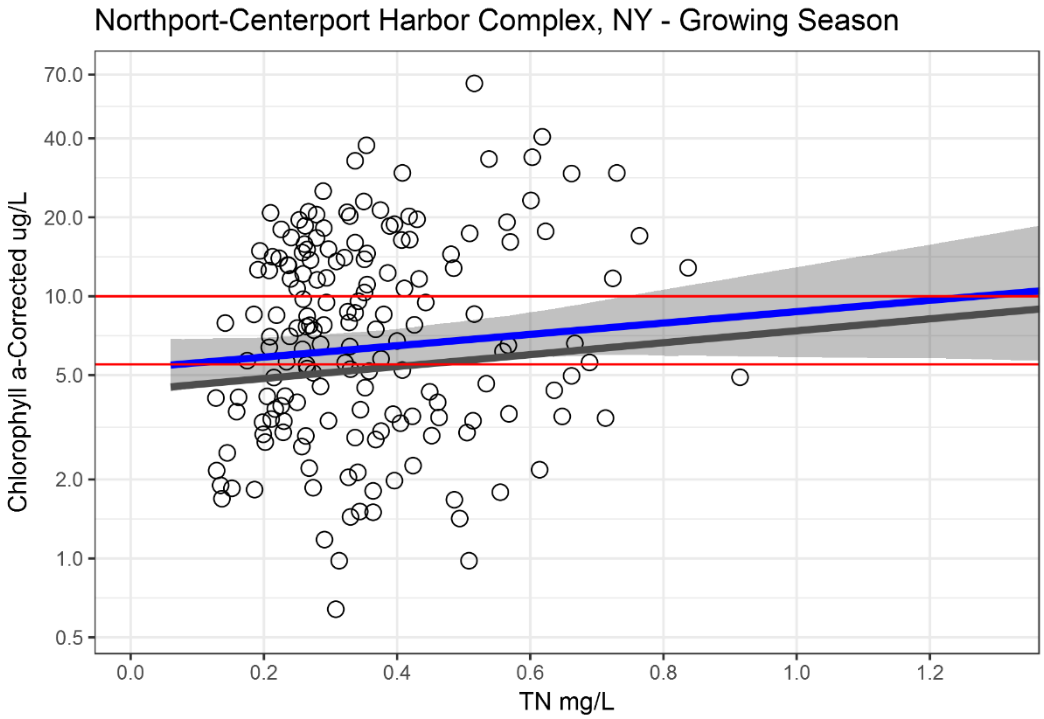

Figure G-18. Chlorophyll vs. Total Nitrogen Relationship for the Northport-Centerport Harbor Complex Watershed, NY (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-9.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoints to protect light levels needed to support seagrasses are 0.07 mg/L and 1.27 mg/L. Respectively, these values are based on the embayment-specific chlorophyll *a* endpoints of 5.5 μg/L (consistent with the value Vaudrey (2008) recommended for LIS) and 10 μg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). These values are further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.10 Port Jefferson Harbor, NY

Figure G-19 shows a map of the Port Jefferson Harbor watershed. Paired data for the embayment included 221 observations across 12 water quality stations within the growing season (April–September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

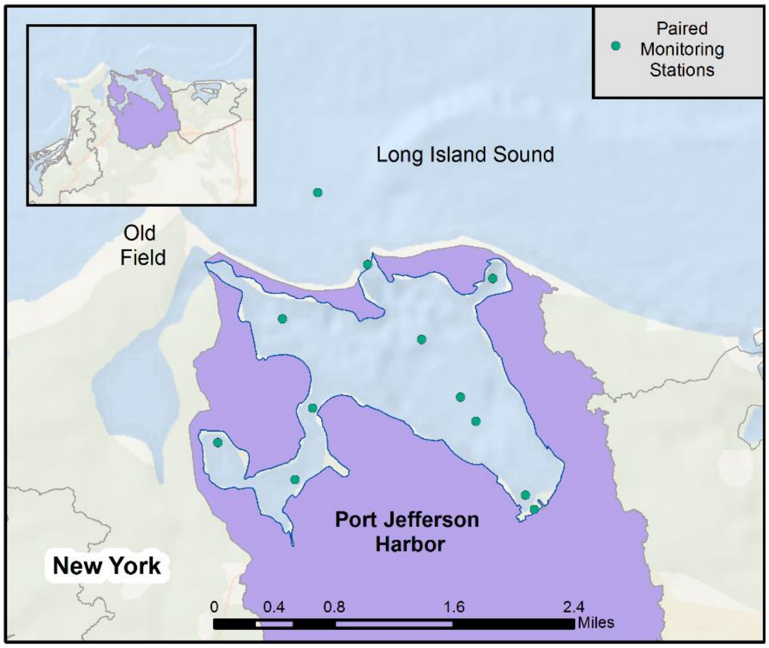

Figure G-19. Port Jefferson Harbor Watershed, NY

Chlorophyll endpoints and TN endpoints for the Port Jefferson Harbor watershed are presented in Table G-10. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-20.

Table G-10. Chlorophyll and TN Endpoints for the Port Jefferson Harbor Watershed, NY

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	5.5	**0.55**[a,b]
		10	**1.69**[a,b]
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).
[b] Calculated using population level model.

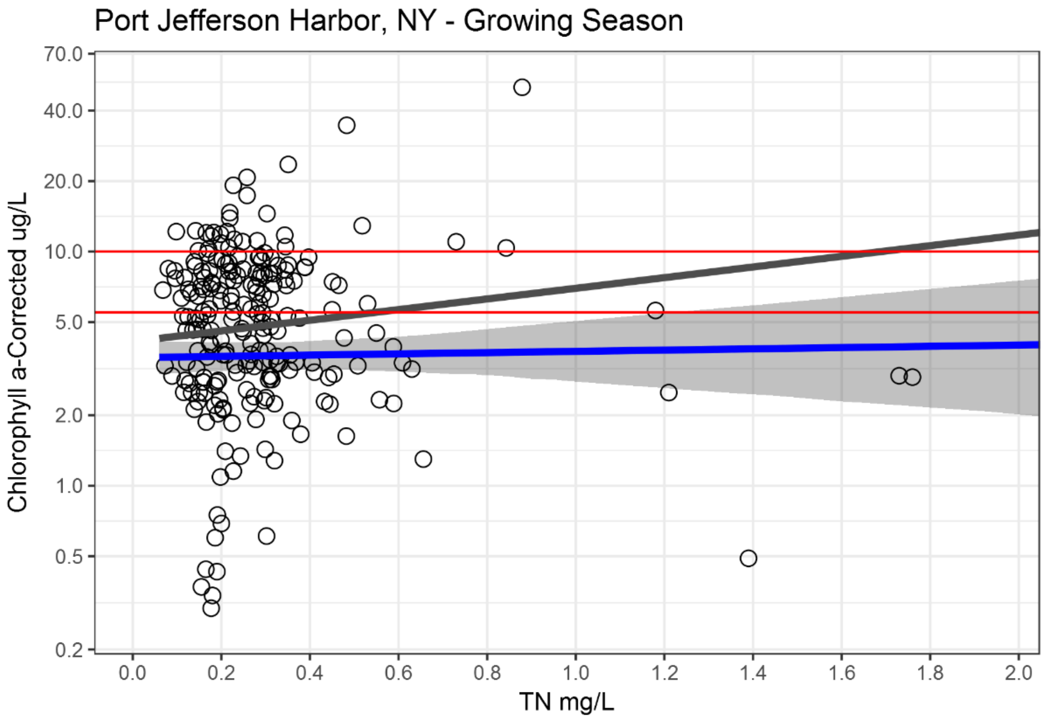

Figure G-20. Chlorophyll vs. Total Nitrogen Relationship for the Port Jefferson Harbor Watershed, NY (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-10.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoints to protect light levels needed to support seagrasses are 0.55 mg/L and 1.69 mg/L. Respectively, these values are based on the embayment-specific chlorophyll *a* endpoints of 5.5 μg/L (consistent with the value Vaudrey (2008) recommended for LIS) and 10 μg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). These values are further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.11 Nissequogue River, NY

Figure G-21 shows a map of the Nissequogue River watershed. Paired data for the embayment included 26 observations across five water quality stations within the growing season (April–September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

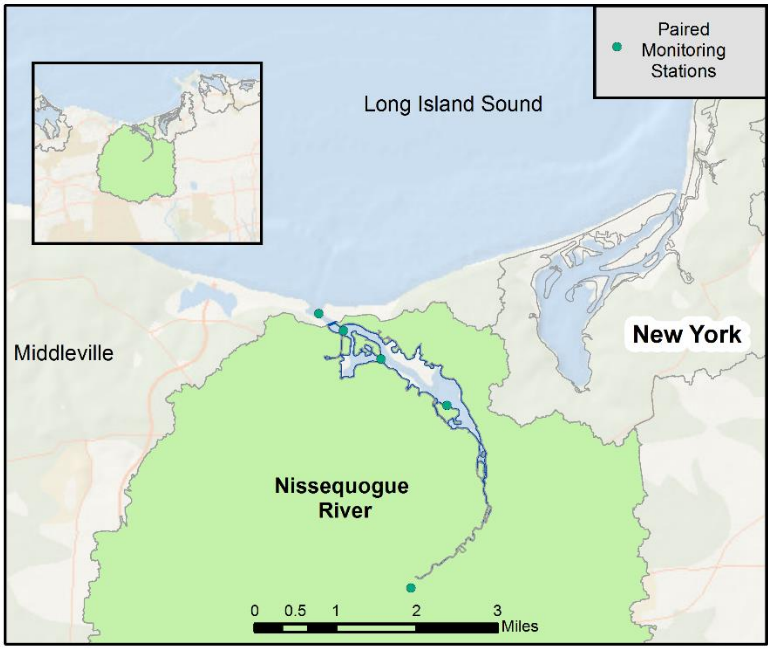

Figure G-21. Nissequogue River Watershed, NY

Chlorophyll endpoints and TN endpoints for the Nissequogue River watershed are presented in Table G-11. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-22.

Table G-11. Chlorophyll and TN Endpoints for the Nissequogue River Watershed, NY

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	5.5	**1.21**[a] (0.06–2.52)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

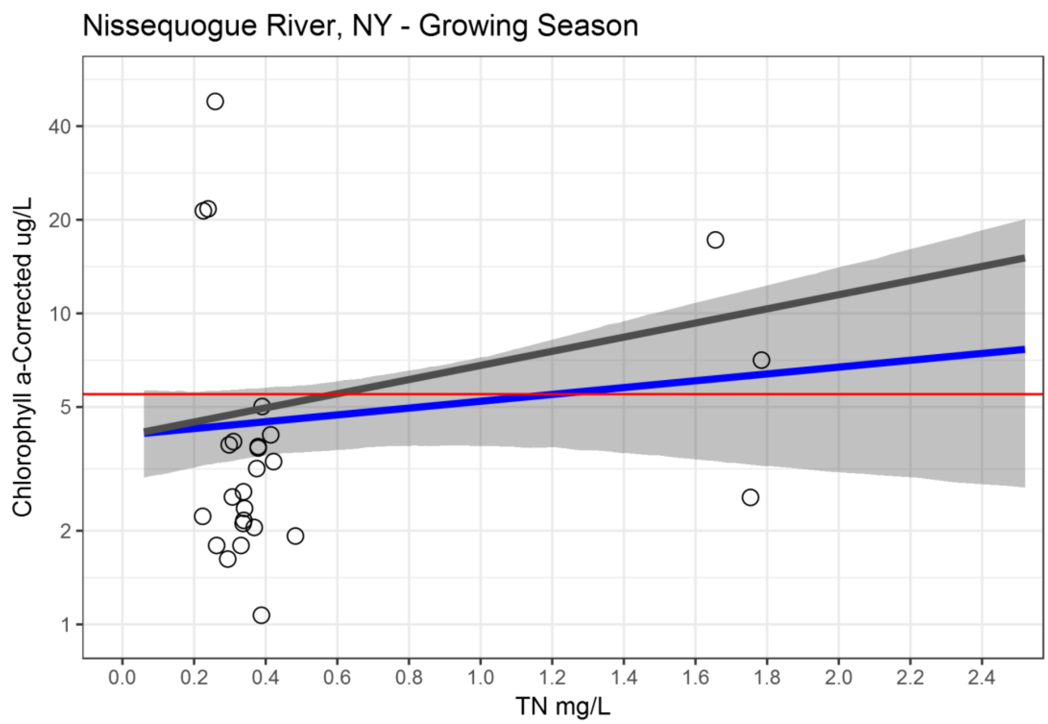

Figure G-22. Chlorophyll vs. Total Nitrogen Relationship for the Nissequogue River Watershed, NY (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-11.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 1.21 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 5.5 µg/L (consistent with the value Vaudrey (2008) recommended for LIS). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.12 Stony Brook Harbor, NY

Figure G-23 shows a map of the Stony Brook Harbor watershed. Paired data for the embayment included 64 observations across seven water quality stations within the growing season (April to September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

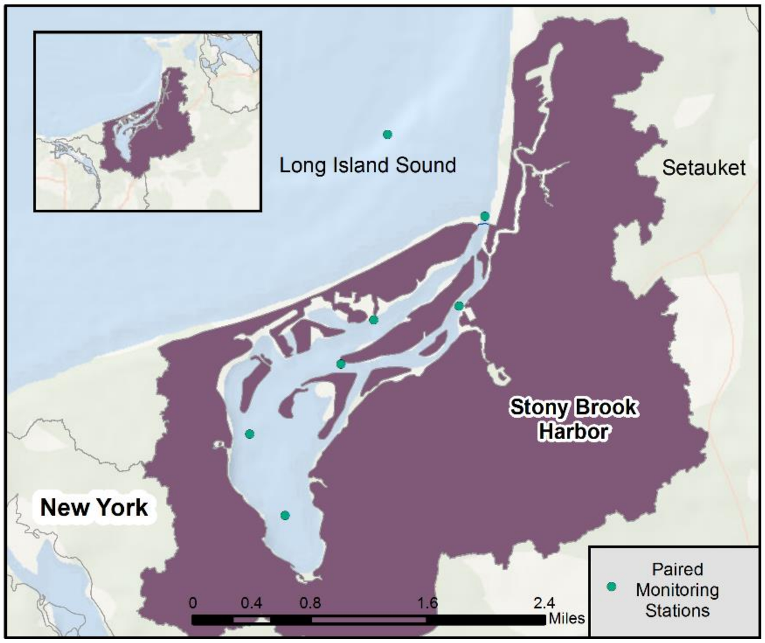

Figure G-23. Stony Brook Harbor Watershed, NY

Chlorophyll endpoints and TN endpoints for the Stony Brook Harbor watershed are presented in Table G-12. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-24.

Table G-12. Chlorophyll and TN Endpoints for the Stony Brook Harbor Watershed, NY

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	5.5	**0.45**[a]
		10	**1.59**[a,b]
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] Calculated using population level model.
[b] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

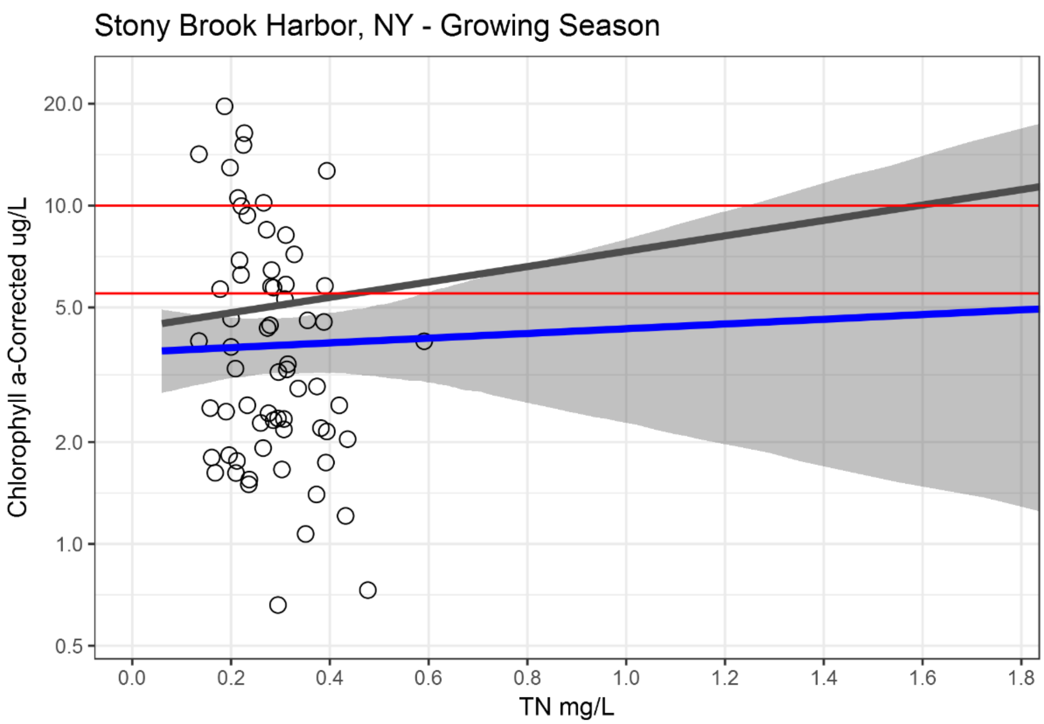

Figure G-24. Chlorophyll vs. Total Nitrogen Relationship for the Stony Brook Harbor Watershed, NY (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-12.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoints to protect light levels needed to support seagrasses are 0.45 mg/L and 1.59 mg/L. Respectively, these values are based on the embayment-specific chlorophyll *a* endpoints of 5.5 μg/L (consistent with the value Vaudrey (2008) recommended for LIS) and 10 μg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). These values are further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.13 Mt. Sinai Harbor, NY

Figure G-25 shows a map of the Mt. Sinai Harbor watershed. Paired data for the embayment included 40 observations across six water quality stations within the growing season (April–September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. In addition to paired data for this embayment, the hierarchical model also incorporates data from all 1,335 embayment observations LIS-wide. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

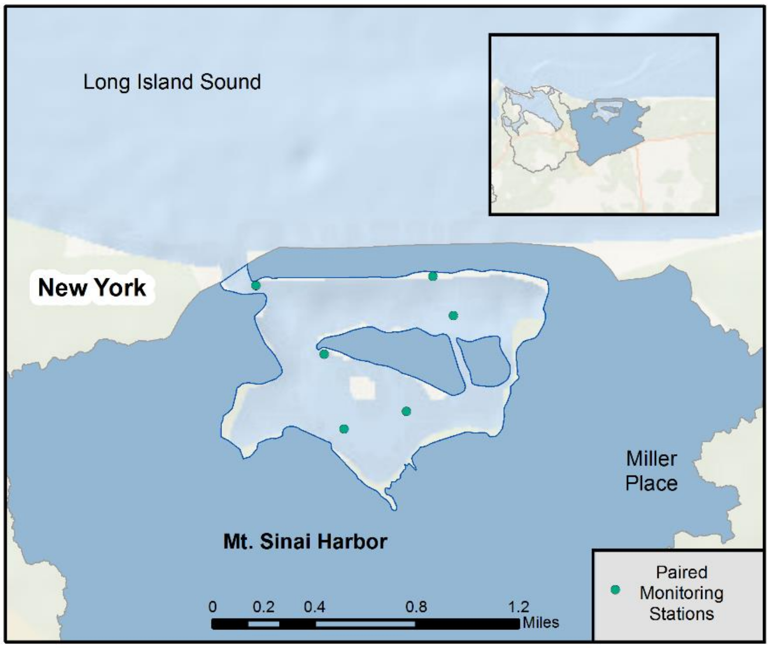

Figure G-25. Mt. Sinai Harbor Watershed, NY

Chlorophyll endpoints and TN endpoints for the Mt. Sinai Harbor watershed are presented in Table G-13. A scatterplot of the chlorophyll versus TN relationship is presented in Figure G-26.

Table G-13. Chlorophyll and TN Endpoints for the Mt. Sinai Harbor Watershed, NY

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	5.5	**1.56**[a] (0.54–2.52)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

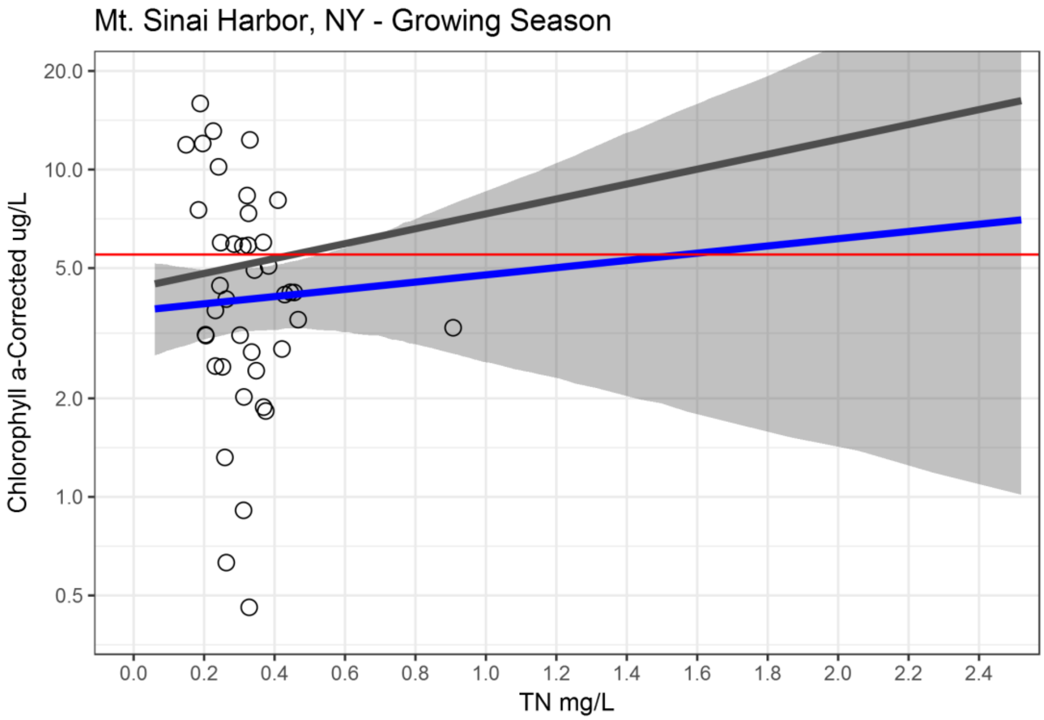

Figure G-26. Chlorophyll vs. Total Nitrogen Relationship for the Mt. Sinai Harbor Watershed, NY (Blue Line is Embayment-Adjusted Fit; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-13.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 1.56 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 5.5 μg/L (consistent with the value Vaudrey (2008) recommended for LIS). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

G.14 Eastern and Western Narrows (Combined), CT and NY

Figure G-27 shows a map of the Eastern and Western Narrows watersheds (combined). Paired data for the open water included 1,157 open water observations across 18 water quality stations within the growing season (April–September). These data were obtained from the water quality data used for analyzing the watershed in Subtask D. As described in Subtask F, parameters used in the model include chlorophyll *a*-corrected, TN, and pH where available. However, stressor-response models were not significant for the open water segments as described in Subtask F.

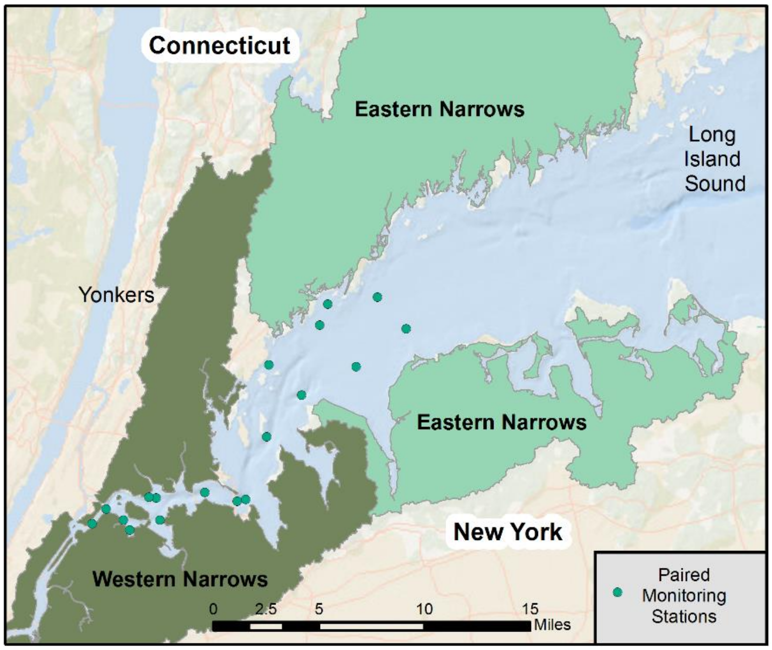

Figure G-27. Eastern and Western Narrows (Combined) Watersheds, CT and NY

Open water TN endpoints for the Eastern and Western Narrows watersheds (combined) are presented in Table G-14.

Table G-14. Chlorophyll and TN Endpoints for the Eastern and Western Narrows (Combined) Watersheds, CT and NY

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Aquatic Life Protection	Literature Review Median Protective of All Endpoints (Range)		**0.40** (0.30–0.60)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.24**

TN Endpoints Discussion

The resulting values for each line of evidence are given in Table G-14.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

G.15 Connecticut River, CT

Figure G-28 shows a map of the Connecticut River where it enters LIS, including the 40% isopleth. The estuarine area of influence of the Connecticut River as described in the Task E memo and indicated on the figure below was the focus of the stressor-response modeling and of the resulting values. No paired data was available for the embayment within the growing season (April–September). Therefore, the global fit using data from all 1,335 embayment observations LIS-wide was used for the stressor-response analysis. As described in Subtask F, parameters used in the hierarchical model include chlorophyll *a*-corrected, TN, pH, and temperature where available.

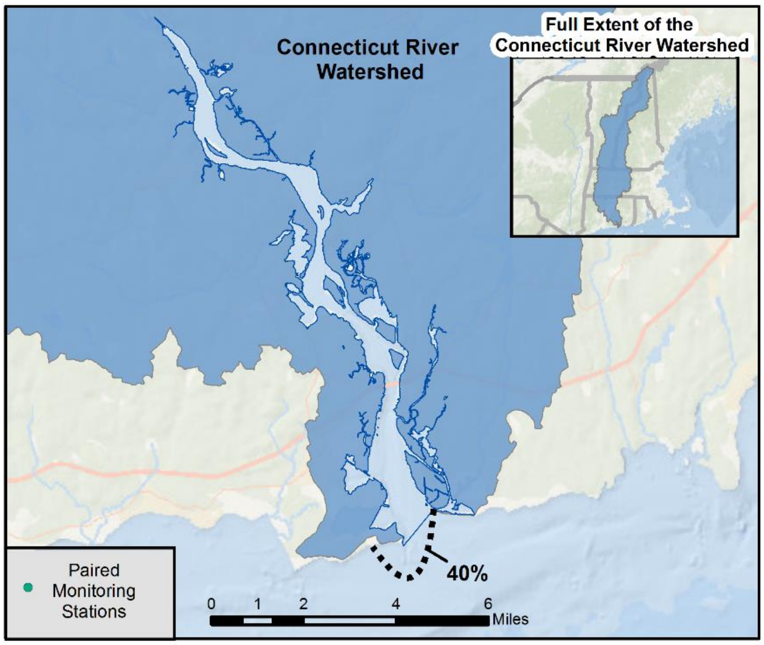

Figure G-28. Connecticut River

Chlorophyll endpoints and TN endpoints for the Connecticut River watershed are presented in

Table G-15. A plot of the chlorophyll versus TN relationship is presented in Figure G-29.

Table G-15. Chlorophyll and TN Endpoints for the Connecticut River Area of Influence

Endpoint Parameter	Endpoint Method	Endpoint Value (µg/L)	TN Endpoint (mg/L)
Chlorophyll *a*-corrected	Stressor–Response Model for Individual Embayments Mean (90th Percent Confidence Interval)	10	**0.87**[a] (0.76–0.97)
Aquatic Life Protection	Literature Review Median Protective of Seagrass Endpoints (Range)		**0.40** (0.30–0.50)
	Distribution-Based Approach – All Embayments 25th Percentile		**0.27**

[a] As per literature review, values exceeding 0.49 mg/L are not considered protective of eelgrass and above 0.60 mg/L are not protective of other endpoints. Values below 0.20 mg/L are considered below background levels (Howes et al. 2006; NHDES 2009).

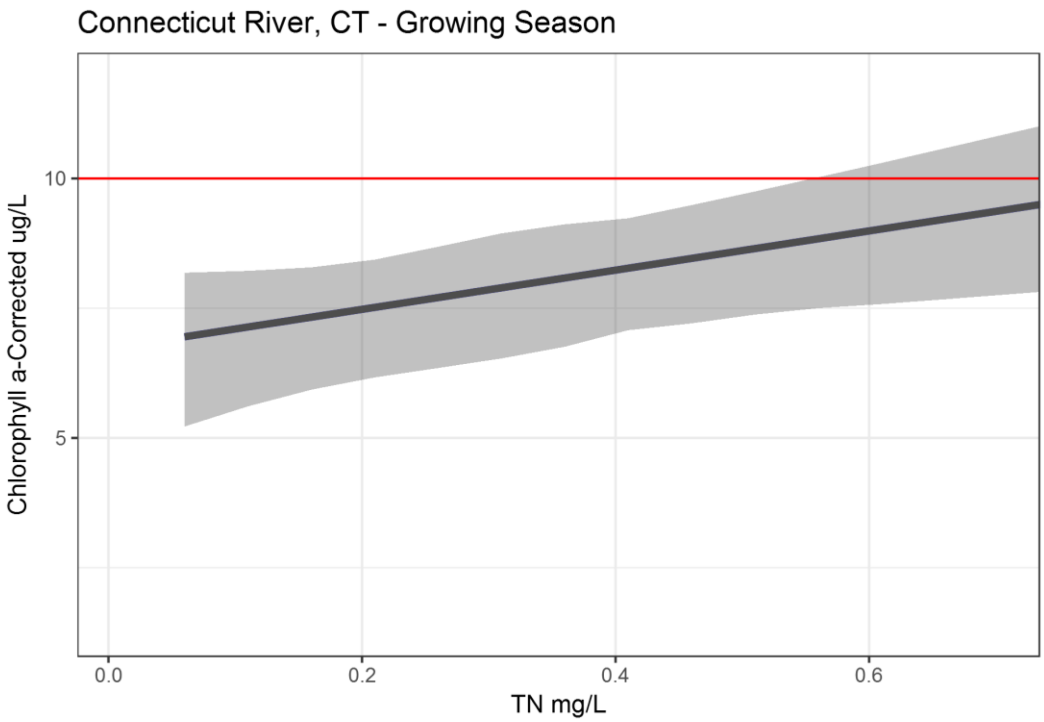

Figure G-29. Chlorophyll vs. Total Nitrogen Relationship for the Connecticut River Area of Influence (There Were No Paired Growing Season Observations Available for this Embayment; Gray Line is LIS Population Fit; Red Lines are Chlorophyll Endpoint Values; Gray Ribbon is 90% Confidence Interval)

TN Endpoints Discussion

The resulting values for each line of evidence are given in

Table G-15.

Literature reviews and distribution-based lines of evidence yield TN values of 0.40 and 0.27 mg/L, respectively (0.40 mg/L is the median literature value, with a minimum of 0.30 mg/L and a maximum of 0.50 mg/L). These represent regionally relevant TN values or concentrations for protecting comparable aquatic life uses (e.g., seagrasses and benthic fauna) based on water quality modeling and ecological research (Howes et al. 2003), as well as on the conditions in water bodies expected to support aquatic life uses including seagrasses and benthic fauna.

For this embayment, the relevant stressor-response TN endpoint to protect light levels needed to support seagrasses is 0.87 mg/L. This value is based on the embayment-specific chlorophyll *a* endpoint of 10 μg/L (consistent with a K_d of 0.7, a value considered on the upper end for protecting light levels necessary for seagrasses in LIS [Vaudrey 2008]). This value is further explained and justified in Subtask F. The embayment stressor-response models often produced TN values that were too low (below most regional background levels and thus not realistic to achieve) or too high (not protective of eelgrass). Instances where this occurred are noted in the embayment endpoint table. EPA plans to revisit the assumptions made during the stressor-response analysis in the next phase of this work.

Subtask G Sources Cited

Howes, B.L., R. Samimy, and B. Dudley. 2003. *Site-Specific Nitrogen Thresholds for Southeastern Massachusetts Embayments: Critical Indicators Interim Report*. Prepared by Massachusetts Estuaries Project for the Massachusetts Department of Environmental Protection. Accessed February 2017. http://yosemite.epa.gov/OA/EAB_WEB_Docket.nsf/Verity%20View/DE93FF445FFADF1285257527005AD4A9/$File/Memorandum%20in%20Opposition%20...89.pdf.

Howes, B.L., S.W. Kelley, J.S. Ramsey, R. Samimy, D. Schlezinger, and E.M. Eichner. 2006. *Linked Watershed-Embayment Model to Determine Critical Nitrogen Loading Thresholds for Nantucket Harbor, Town of Nantucket, Nantucket Island, Massachusetts*. SMAST/DEP Massachusetts Estuaries Project. Massachusetts Department of Environmental Protection, Boston, MA. Accessed March 2018. http://www.oceanscience.net/estuaries/report/Nantucket/Nantucket_Hbr_MEP_Final.pdf.

NHDES. 2009. *Numeric Nutrient Criteria for the Great Bay Estuary*. New Hampshire Department of Environmental Services, Concord, New Hampshire. Accessed March 2018. https://www.des.nh.gov/organization/divisions/water/wmb/wqs/documents/20090610_estuary_criteria.pdf.

Vaudrey, J.M.P. 2008. *Establishing Restoration Objectives for Eelgrass in Long Island Sound. Part I: Review of the Seagrass Literature Relevant to Long Island Sound.* Final Grant Report to the Connecticut Department of Environmental Protection, Bureau of Water Protection and Land Reuse and the U.S. Environmental Protection Agency.

Vaudrey, J.M.P, C. Yarish, J.K. Kim, C. Pickerell, L. Brousseau, J. Eddings, and M. Sautkulis. 2016. *Connecticut Sea Grant Project Report: Comparative Analysis and Model Development for Determining the Susceptibility to Eutrophication of Long Island Sound Embayments*. Project number R/CE-34-CTNY. 46 p.