

Identification of Algae-Nutrient Relationships

David A. Chin*

Department of Civil Engineering, University of Miami, Florida, USA

Abstract: Fundamental relationships are presented showing that the probability distribution of chlorophyll *a* (Chla) depends on both the distribution of Chla conditioned on TN and TP, and the distribution of TN conditioned on TP. It is further shown that simplified forms of these distributions can be identified using the non-parametric Spearman correlations between Chla, TN, and TP. Based on these fundamental relationships, a novel approach is presented for estimating Chla concentrations that do not depend on a-priori assumptions of an analytic relationship between Chla, TN, and TP. The proposed approach is demonstrated by application to six subareas within an impaired segment of the Halifax River in Florida. Results from these analyses show that the relationship between Chla, TN, and TP exhibits significant spatial and temporal variability over length scales on the order of 1 mile, and that using the Redfield ratio as a basis for identifying the controlling nutrient is not reliable.

Keywords: Chlorophyll *a*, Impaired Water, Nitrogen, Phosphorus, TMDL.

1. INTRODUCTION

Excessive algae concentrations in natural waters are generally undesirable due to their negative effects on water quality, biological diversity, engineered systems, and aesthetics. The concentration of algae is commonly measured as chlorophyll *a* (Chla). Algal growth in many natural water bodies is primarily influenced by dissolved species of nitrogen (N) and/or phosphorus (P), with other water-quality parameters such as pH, conductivity, levels of organic contamination, sediments, and pesticides influencing algal growth to a lesser extent. The aqueous concentrations of N and P are commonly represented by the concentrations of total nitrogen (TN) and total phosphorus (TP), respectively. It has long been accepted that the relationship between Chla, TN, and TP depends on the trophic status of the water body [1, 2]. In fresh water, Chla concentrations tend to be controlled primarily by TP under oligotrophic and mesotrophic conditions (where $TP < 30 \mu\text{g/L}$) and controlled primarily by TN under hypereutrophic conditions ($TP > 1000 \mu\text{g/L}$). In between these trophic states Chla concentrations can be expected to depend on both TN and TP. The form of the Chla-TN-TP relationship is commonly referenced to the TN/TP ratio, and a common (mis-)perception is that a critical value of TN/TP exists such that when TN/TP in a water body is less than the critical value Chla depends only on TN, and when TN/TP is greater than the critical value Chla depends only on TP. The most commonly used critical TN/TP value is the Redfield ratio, first proposed by Redfield (1934) [3] who noted that plankton cells have an average atomic ratio of $\text{C}_{106}\text{N}_{16}\text{P}$, which indicates that 16 moles of N are required for each mole of P. On a mass basis, the Redfield ratio corresponds to

TN/TP equal to 7.2. The fallacy of using the Redfield ratio as a basis for identifying the critical nutrient is that there are usually many algal species present with a range of molar ratios of N and P in their biomass. Observations generally indicate that the critical TN/TP ratio in a water body is significantly related to the trophic state, with high values of TN/TP (~ 40) characteristic of oligotrophic waters, intermediate values of TN/TP (~ 20) characteristic of mesotrophic waters, and low values of TN/TP (~ 5) characteristic of eutrophic and hypereutrophic waters.

2. ANALYSIS OF MEASUREMENTS

Measurements of Chla, TN, and TP are sometimes averaged prior to analysis, with growing-season average and annual average being fairly common. Various approaches have been used to analyze the relationship between Chla, TN, and TP and these approaches can be broadly categorized as empirical or mechanistic. Empirical approaches typically rely on linear regression between variables, whereas mechanistic approaches are generally based on simplified physical, chemical, and/or biological relationships between variables along with various assumed process parameters. The major limitation of empirical approaches is the assumption of linearity in the relationship between variables. Fitting linear functions to nonlinear Chla-TN-TP relationships can be a significant source of error [2], and empirical formulations that accommodate nonlinearity have been relatively few. An example of a quasi-nonlinear analysis was presented by Seip (1994) [4] who applied piecewise linear regression by separating the data into TP categories that roughly corresponded to different trophic states.

A variety of factors besides nutrient concentrations can affect Chla concentrations in aquatic environments. These factors include: temperature, light, salinity, water hardness, and the presence of benthic algae [5, 6]. In spite of the variety of

*Address correspondence to this author at the Department of Civil Engineering, University of Miami, Florida, USA;
E-mail: dchin@miami.edu

potentially influential factors, it is commonly reported that TP is the limiting nutrient in freshwater systems and TN is the limiting nutrient in saltwater systems. However, this rule is not universal since there have been some reports of TP regulating Chla concentrations in both freshwater and marine aquatic environments [7]. Also of importance in developing Chla-TN-TP relationships is the so-called vernal bloom hypothesis [8, 9], which suggests that spring (vernal) algal blooms are commonly caused primarily by light, and are not significantly influenced by aqueous concentrations of TN and TP.

A complicating factor that limits the ability to derive useful Chla-TN-TP relationships from measurements is the relatively large spatial variability of Chla, TN, and TP in aquatic systems, a condition that has led some investigators to conclude that the content of a sampling bottle might be a poor reflection of the actual conditions within a water body [10]. This reality leads to significant uncertainty in any parameter or statistic that is derived from highly variable measurements. Useful guidance on estimating the mean value of sampled data from normally distributed population is [11]

$$L = \frac{t \cdot CV}{\sqrt{n-1}} \quad (1)$$

where L is the fractional error of the estimated mean within the assumed confidence interval, t is the Student's t variable corresponding to the assumed confidence interval, CV is the coefficient of variation of the population distribution (= mean/standard deviation), and n is the number of samples. For example, if 5 samples are collected in an environment in which the variable has a coefficient of variation of 0.33 and a 95% confidence interval is assumed, then $t = 1.96$, $CV = 0.33$, $n = 5$, and Equation 1 gives $L \approx 0.33$, which indicates that there is potentially a 33% error in the estimated value of the mean. It is apparent that the reliability of analytic or mechanistic models that are calibrated to measured data can be severely limited by having only a small number of measurements.

Concentrations of TN and TP in aquatic systems are commonly related to each other [10], a condition that is attributed to the fact that these nutrients typically have common sources and terrestrial transport mechanisms. For example, primary sources of nutrients in a water body might be agricultural runoff, urban drainage, and/or direct wastewater discharges, all of which have characteristic TN/TP signatures. Increased TP concentrations typically correspond to increased TN concentrations and vice versa, however, analyses of data from several lakes around the world indicate that this relationship is usually nonlinear [1]. The relationship between TN and TP is further complicated by the fact that, aside from TN and TP loading from external sources, TN and TP concentrations are commonly influenced by *in situ* processes such as sedimentation, nitrification, denitrification, and nitrogen fixation. As a consequence, the relationship between TN and TP typically varies with the trophic state of the water body.

3. FUNDAMENTALS

The relationship of interest in water bodies that are impaired by excessive algal growth is the relationship between

Chla, TN and TP. This relationship is usually the basis for quantifying nutrient reductions that are necessary to decrease Chla concentrations to acceptable levels. Since Chla, TN, and TP are all stochastic variables, these variables must necessarily be specified by their probability distributions and their relationship described in probabilistic terms. The general probabilistic relationship between these variables and a data analysis protocol that preserves the stochastic nature of these variables are described below.

3.1. Theory

The probability distribution of Chla, denoted by $p(\text{Chla}|\alpha)$, is related to the probability distribution of TN and TP by the relation

$$p(\text{Chla}|\alpha) = \int_0^\infty f(\text{Chla}|\text{TN},\text{TP},\alpha)g(\text{TN}|\text{TP})h(\text{TP})d\text{TP} \quad (2)$$

where α represents the variables other than TN and TP that affect the concentration of Chla; $f(\text{Chla}|\text{TN},\text{TP},\alpha)$ is the conditional probability distribution of Chla for given values of TN, TP, and α ; $g(\text{TN}|\text{TP})$ is the conditional probability distribution of TN for given values of TP; and $h(\text{TP})$ is the probability distribution of TP. In cases where only TN and TP control the growth of algae in a water body, Equation 2 can be expressed as

$$p(\text{Chla}) = \int_0^\infty f(\text{Chla}|\text{TN},\text{TP})g(\text{TN}|\text{TP})h(\text{TP})d\text{TP} \quad (3)$$

It is apparent from Equation 3 that the probability distribution of Chla is influenced by two independent processes: (1) the relationship between Chla and TN and TP, which is denoted by $f(\text{Chla}|\text{TN},\text{TP})$; and (2) the relationship between TN and TP, which denoted by $g(\text{TN}|\text{TP})$. Further simplifications can be made to Equation 3, depending on whether there is a single controlling nutrient or no controlling nutrient at all, and whether the variables TN and TP are independent. Assuming that measurements of Chla, TN and TP are available, then the appropriate functional forms in Equation 3 can be determined by calculating the following Spearman correlation coefficients: $r(\text{Chla},\text{TP})$, $r(\text{Chla},\text{TN})$, and $r(\text{TN},\text{TP})$. Since the Spearman correlation coefficients are nonparametric, they do not require that the relationships between the variables be linear. The functional forms of $f(\text{Chla}|\text{TN},\text{TP})$ and $g(\text{TN}|\text{TP})$ that can be inferred from the Spearman correlations are given in Table 1.

The determination of whether the Spearman correlation coefficient, r , is significantly nonzero can be done using the following variable transformation [12],

$$t = \sqrt{\frac{n-2}{1-r^2}} \quad (4)$$

where t has a Student's t distribution with $n - 2$ degrees of freedom under the null hypothesis that $r = 0$, and n is the number of measurements. Typically, t values outside of the 90% confidence interval of t are taken to indicate that the correlation is (significantly) nonzero. Identification of the appropriate functional forms of $f(c|\text{TN},\text{TP})$ and $g(\text{TN}|\text{TP})$ have significant practical consequences. For example, if $f(c|\text{TN},\text{TP})$ reduces to $f(c)$ then Chla concentrations cannot be controlled by adjusting TN and TP concentrations and other controlling variables must be sought, also, if $g(\text{TN}|\text{TP})$ reduc-

Table 1. Chla-nutrient relationships under various conditions.

Function	Correlations	Reduced Function
$f(c TN,TP)$	$r(\text{Chla},TP) = 0$ and $r(\text{Chla},TN) = 0$	$f(c)$
	$r(\text{Chla},TP) \neq 0$ and $r(\text{Chla},TN) = 0$	$f(c TP)$
	$r(\text{Chla},TP) = 0$ and $r(\text{Chla},TN) \neq 0$	$f(c TN)$
	$r(\text{Chla},TP) \neq 0$ and $r(\text{Chla},TN) \neq 0$	$f(c TN,TP)$
$g(TN TP)$	$r(TN,TP) = 0$	$g(TN)$
	$r(TN,TP) \neq 0$	$g(TN TP)$

es to $g(TN)$ then TN and TP levels can likely be adjusted independently (by source controls), otherwise it is unlikely that TN and TP can be adjusted independently and that adjustment of either TN or TP will influence the Chla concentrations.

3.2. Data Analysis

The minimum data needed to relate Chla to nutrient levels are synoptic measurements of Chla, TP, and TN in seasons where excessive algal growth is a concern. The distribution of Chla concentrations, c , can be expressed in the following discrete form,

$$S(c) = \sum_{i=1}^I \sum_{j=1}^J S(c|[TN]_i[TP]_j) = \sum_{i=1}^I \sum_{j=1}^J \{c_1, c_2, \dots, c_{n_{i,j}}\}_{i,j} \quad (5)$$

where $S(c)$ is the set of all measured Chla concentrations; $S(c|[TN]_i[TP]_j)$ is the set of measured Chla concentrations when TN and TP are in their i th and j th quantiles, denoted by $[TN]_i$ and $[TP]_j$, respectively; $\{c_1, c_2, \dots, c_{n_{i,j}}\}_{i,j}$ is the set of measured Chla concentrations when TN and TP are in $[TN]_i$ and $[TP]_j$, respectively; and $n_{i,j}$ is the number of measurements in $[TN]_i$ and $[TP]_j$. Representing the algal concentrations as discrete values provides the opportunity to use nonparametric ranking to estimate the exceedance probability of any specified value of c , provided that there are an adequate number of data points.

If the nutrient concentrations in the water body are changed, such as by changing the nutrient loading, then it can be expected that the Chla concentrations might also change in response. Denoting the changed concentrations by c^* , then the set of changed concentrations, $S(c^*)$ can be expressed in the form

$$S(c^*) = \sum_{i=1}^I \sum_{j=1}^J S(c^*|[TN]_i[TP]_j) = \sum_{i=1}^I \sum_{j=1}^J \{c_1^*, c_2^*, \dots, c_{n_{i,j}^*}\}_{i,j} \quad (6)$$

where $n_{i,j}^*$ is the changed number of measurements in the original quantiles $[TN]_i$ and $[TP]_j$. It is possible to estimate changed concentrations, c^* , from the measured concentrations, c . Consider first the case in which TN and TP are not correlated, which means that TN and TP can be independently changed. If α_N and α_P are the reduction factors of TN and TP, respectively, then each measurement of TN and TP in the TN versus TP plot (i.e., in the TN-TP plane)

is transformed to a new location ($\alpha_N TN$, $\alpha_P TP$). After transforming all of the measurement points, the number of points in each original (i,j) quantile cell can then be counted and denoted by $n_{i,j}^*$. Assuming that the functional relationship between Chla, TN, and TP is stationary, and noting that the actual measurements of Chla in cell (i,j) are $\{c_1, c_2, \dots, c_{n_{i,j}}\}_{i,j}$, then the Chla concentrations corresponding to the new points in cell (i,j) can be obtained by resampling $n_{i,j}^*$ values from the set $\{c_1, c_2, \dots, c_{n_{i,j}}\}_{i,j}$; this resampled set is $\{c_1^*, c_2^*, \dots, c_{n_{i,j}^*}\}_{i,j}$.

In cases where TN and TP are correlated, reductions in TN and TP cannot be specified independently, and a reduction in either TN or TP must be specified. If the fractional reduction in TP is specified as α_P , then each TP measurement is reduced to $\alpha_P TP$ and the new number of points in each original TP quantile, $[TP]_j$, can be calculated. If the number of transformed points in $[TP]_j$ is n_j^o , then the new number of points in cell (i,j), denoted by $n_{i,j}^*$ can then be estimated as

$$n_{i,j}^* = \left\| \frac{n_{i,j}}{\sum_{i=1}^I n_{i,j}} n_j^o \right\| \quad (7)$$

where $\| \cdot \|$ represents a nearest integer operation, and $n(i,j)$ is the original distribution of points within the quantile cells in the TN-TP plane. The desirable feature of Equation 7 is that it preserves the original TN-TP relationship while applying a reduction factor to all of the original TP measurements. The changed concentrations in cell (i,j) can then be obtained by resampling $n_{i,j}^*$ values from the set $\{c_1, c_2, \dots, c_{n_{i,j}}\}_{i,j}$, and this resampled set is equal to $\{c_1^*, c_2^*, \dots, c_{n_{i,j}^*}\}_{i,j}$. This

procedure can also be applied to determine the effect of reducing TN instead of TP; in which case the roles of TN and TP are reversed in the analysis. It should be noted that both TP reduction and the TN reduction should be investigated, since even though TN and TP are correlated, reductions in Chla resulting from a given fractional reduction in TP will likely be different from the reductions in Chla resulting from the same fractional reduction in TN.

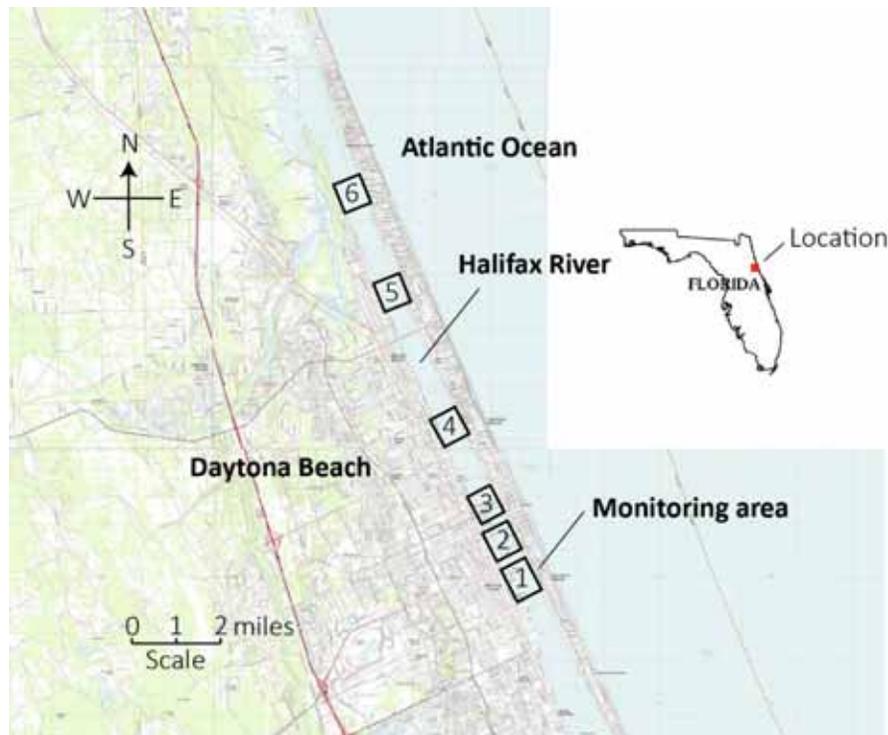


Fig. (1). Monitoring areas in the Halifax River, Florida.

4. DEMONSTRATION

The Halifax River is a 37-km (23-mi) long tidal estuary located near the east coast of northern Florida as shown in Fig. (1). The state of Florida has designated the Halifax River as being impaired due to excessive nutrients, which is manifested by excessive algae growth. As a consequence of being designated as impaired, federal regulations require that a total maximum daily load (TMDL) of nutrients be established that is consistent with the water body meeting its regulatory water-quality criteria for Chla. Synoptic measurements of Chla, TN, and TP concentrations in the impaired segment of the Halifax River have been reported by Magley (2013) [13] for the years between 1968 and 2011. These data were collected at several monitoring stations within the impaired segment, and these stations can be grouped within monitoring areas that each extend approximately 1.6 km (1 mi) along the river as shown in Fig. (1). The monitoring stations within these areas are listed in Table 2 by their STORET identification numbers.

To account for seasonal variations in potentially important factors such as sunlight, water temperature, and anthropogenic activities in the surrounding watershed, separate analyses were done for each of four seasons designated by the three-month periods of January-March, April-June, July-September, and October-December. In order to determine the trophic state in the river, the mean TN/TP ratio was calculated for each season in each monitoring area, and these results are shown as TN/TP in Table 2. It is apparent from these results that TN/TP falls within the relatively narrow range of

5.1–8.1 for all monitoring areas, and based on conventional classifications [14] the Halifax River is eutrophic year round.

To assess the relationship between Chla, TN, and TP within each monitoring area and season, the Spearman correlation coefficients $r(\text{Chla}, \text{TN})$ and $r(\text{Chla}, \text{TP})$ were calculated and the reduced functional forms for $f(c|\text{TN}, \text{TP})$ were determined in accordance with Table 1. Similarly, to assess the relationship between TN and TP the Spearman correlation coefficient $r(\text{TN}, \text{TP})$ was calculated and the reduced forms for $g(\text{TN}/\text{TP})$ were also determined in accordance with Table 1. The results of these analyses are shown in Table 3, where the monitoring area, season, Chla function, TN function, and the numbers of measurements used in the analyses are given in Columns 1–5, respectively. Notable features of these results are: (1) the functional relationship between Chla, TN, and TP varies between seasons; (2) in 5 of the 6 monitoring areas investigated, Chla is independent of TN and TP in January-March and October-December seasons; (3) in cases where Chla is related to nutrient concentrations, dependence is on TN and TP, or TP alone, and in no case was there dependence on TN alone; and (4) the functional relationship between TN and TP varies between seasons, with independent fluctuations of TN and TP occurring in most monitoring areas during most seasons. The dependence of Chla on both TN and TP during the growing seasons (April-September) in most of the monitoring areas is particularly interesting, since the mean TN/TP ratios in most of the monitoring areas (i.e., all except Area 6) are below the Redfield ratio of 7.2, which according to the limiting-nutrient

Table 2. Monitoring stations and average TN/TP ratios within designated areas.

Area	Stations*	TN/TP
1	21FLVEMD HL12, 21FLCEN 27010121, 21FLFMRIFLEALT158, 21FLCEN 27010119, 21FLSJM27010037, 21FLA 27010037, 21FLA27010038, 21FLVEMDHL11, 21FLVEMDVC-012, 21FLA27010408, 21FLA27010039, 21FLVEMDVC-HL11A, 21FLVEMDVC-011	5.3-6.5
2	21FLCEN 27010120, 21FLFMRIFLEALT2005158, 21FLVEMDHL10, 21FLVEMDVC-010, 21FLCEN 27010946, 21FLA27010946, 21FLSJMHR92DB	5.1-6.2
3	21FLVEMDHL09, 21FLVEMDVC-009, 21FLA27010950, 21FLCEN 27010945, 21FLA27010945, 21FLCEN 27010950, 21FLFMRIFLEALT200107	5.4-6.9
4	21FLA27010944, 21FLVEMDVC-008, 21FLVEMDHL08	5.1-6.4
5	21FLA27010404, 21FLA27010402, 21FLA27010031, 21FLA27010032, 21FLVEMDVC-006, 21FLVEMDHL06, 21FLA27010403	5.7-7.1
6	21FLA27010940, 21FLVEMDVC-005, 21FLA27010831, 21FLVEMDHL05	6.0-8.1

*STORET identification numbers.

Table 3. Monitoring stations within designated areas.

Col 1	Col 2	Col 3	Col 4	Col 5	Col 6	Col 7	Col 8	Col 9
					90-Percentile Chla ($\mu\text{g/L}$) for (α_N, α_P)			
Area	Season	Chla(TN,TP)	TN(TP)	N	(1,1)	(1,0.8)	(0.8,1)	(0.8,0.8)
1	1	$f(c TP)$	$g(TN TP)$	84	13.4	11.0	13.1	-
	2	$f(c TP)$	$g(TN TP)$	110	15.0	13.9	13.9	-
	3	$f(c TP)$	$g(TN TP)$	101	18.0	13.6	16.4	-
	4	$f(c TN,TP)$	$g(TN)$	24	13.2	9.5	17.7	11.1
2	1	$f(c)$	$g(TN TP)$	19	16.8	16.0	14.0	-
	2	$f(c TN,TP)$	$g(TN)$	43	17.9	17.8	16.2	14.2
	3	$f(c)$	$g(TN)$	25	16.3	13.8	14.7	13.3
	4	$f(c)$	$g(TN)$	7	10.6*	-	-	-
3	1	$f(c)$	$g(TN)$	29	17.0	12.7	17.4	11.9
	2	$f(c TN,TP)$	$g(TN)$	45	22.3	17.9	16.6	17.7

Table 3. contd..

Col 1	Col 2	Col 3	Col 4	Col 5	Col 6	Col 7	Col 8	Col 9
	3	$f(c TN,TP)$	$g(TN)$	30	19.9	17.6	16.3	19.0
	4	$f(c)$	$g(TN)$	4	19.4*	-	-	-
4	1	$f(c)$	$g(TN)$	23	17.6	16.7	16.7	16.7
	2	$f(c TN,TP)$	$g(TN TP)$	32	18.9	17.5	18.4	-
	3	$f(c TN,TP)$	$g(TN)$	26	18.0	17.3	17.6	17.6
	4	$f(c)$	$g(TN)$	6	8.7*	-	-	-
5	1	$f(c)$	$g(TN)$	29	12.9	11.2	12.5	11.7
	2	$f(c TN,TP)$	$g(TN)$	38	11.8	12.9	11.8	12.5
	3	$f(c TN,TP)$	$g(TN TP)$	27	22.3	17.8	14.3	14.0
	4	$f(c)$	$g(TN)$	7	30.3*	-	-	-
6	1	$f(c)$	$g(TN)$	26	12.9	12.7	12.1	7.7
	2	$f(c TN,TP)$	$g(TN)$	41	17.7	9.6	16.7	8.6
	3	$f(c TN,TP)$	$g(TN)$	30	15.3	11.5	13.5	11.1
	4	$f(c)$	$g(TN)$	6	23.2*	-	-	-
All	1	$f(c TP)$	$g(TN TP)$	210	15.3	13.0	14.0	-
	2	$f(c TN,TP)$	$g(TN TP)$	309	16.8	14.9	14.7	-
	3	$f(c TN,TP)$	$g(TN TP)$	239	18.2	14.9	16.2	-
	4	$f(c TN,TP)$	$g(TN)$	54	14.0	14.7	16.2	13.1

*Estimate based on < 10 measurement

hypothesis would indicate TN control of Chla levels. It is apparent that the Redfield ratio cannot be used as a strict guideline for determining the controlling nutrient. For regulatory purposes, the segment of the Halifax River that includes all of the monitoring areas is designated as a single impaired water body, and hence the state of Florida lumps all measurements together for analysis. To assess the effect of lumping these data, the functional relationships derived for the individual monitoring areas were also derived for the lumped data and the results are shown in Table 3 under the area designation “All”. It is apparent from these results that the functional relationships derived from the lumped data can deviate significantly from the functional relationships derived from data within the individual monitoring areas, therefore caution should be used in basing local remediation strategies on lumped data. For example, the lumped data show Chla to be dependent on TN and/or TP during all seasons, whereas within individual monitoring areas there are some seasons where Chla is independent of TN and TP. In addition, the lumped data show TN and TP concentrations to be dependent in 3 of the 4 seasons, whereas the data within most monitoring areas show that TN and TP levels fluctuate independently.

The 90-percentile Chla values were calculated for each monitoring area by first ranking the Chla measurements from

lowest to highest and then assigning a percentile value of $(i - 0.5)/n \times 100$, where i is the rank of the datum and n is the number of data points. For the actual measurements, the reduction factors are $\alpha_N = 1$ and $\alpha_P = 1$, and the 90-percentile values, c_{90} , for each monitoring area and season are shown in Column 6 of Table 3. These results show that the maximum c_{90} within each monitoring area occurs in either Season 2 (April–June) or Season 3 (July–September), and the maximum c_{90} varies in the range of 17.7–22.3 $\mu\text{g/L}$ over all of the monitoring areas. Estimation of c_{90} is of limited precision during Season 4 in all monitoring areas except Area 1 due to the small amount of data (< 10 measurements) collected during this season. It would normally be expected that remediation efforts would focus on the season with the maximum c_{90} , and the seasons with maximum c_{90} are shown in bold in Table 3. If the data were lumped (Station “All”), Table 3 shows that Season 3 (July–September) would be identified as the critical season with a c_{90} value of 18.3 $\mu\text{g/L}$. It is noteworthy that acting in local areas on the basis of the lumped-data analysis would be of limited effectiveness, since the critical season in some monitoring areas is Season 2 rather than Season 3, notably Areas 2, 3, 4, and 6.

The effectiveness of reducing nutrient levels on Chla was determined by considering the following values of the

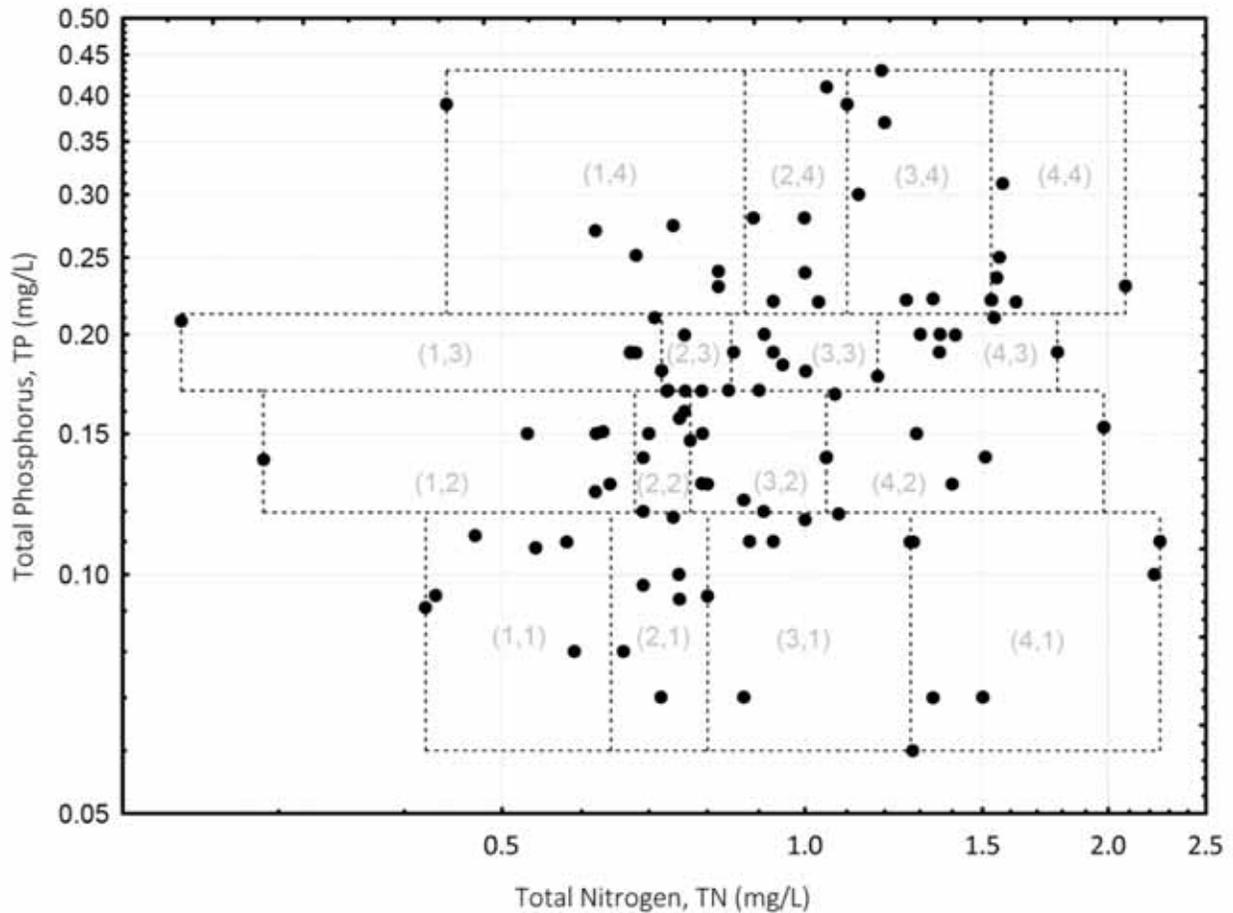


Fig. (2). Percentile-segregated TN and TP measurements in area 1.

reduction factors (α_N, α_P) : (1,0.8), (0.8,1), and (0.8,0.8). The first of these factors represents a TP reduction of 20% with no reduction in TN, the second factor represents a TN reduction of 20% and no reduction in TP, and the third factor represents a 20% reduction in both TN and TP, which can only be specified a-priori if TN and TP are independent. Using the data-adjustment protocol described in the previous section, the c_{90} values resulting from the prescribed nutrient reductions are shown in Table 3. Of practical interest is the identification of which nutrient reduction would be more effective in the critical season, a reduction in TN or a comparable reduction in TP. Taking 20% as the benchmark reduction, the nutrient that produces the lower c_{90} can be identified as the “critical nutrient”, and the critical nutrient reductions within each monitoring area are shown in bold in Table 3. These results indicate that the critical nutrient varies between monitoring areas, with TP being the critical nutrient in Areas 1 and 6, and TN being the critical nutrient in Areas 2–5. Analysis of the lumped data indicate that TP is the critical nutrient. It is apparent that applying the critical-nutrient result from the lumped data analysis to the individual monitoring areas would be of limited effectiveness.

It should be noted that any spurious relationships in the measured data will also be reflected in the analytical

results. For example, if the measured data show that Chla decreases with increasing TN, this will be reflected in the resampled results, as it should be. An illustration of such a case is in Area 4 during Season 4, where a 20% reduction in TN leads to an increase in Chla concentrations. This result is consistent with the reported measurements.

5. NUMERICAL ASPECTS

To illustrate some of the numerical aspects of the data-analysis protocol, consider the data measurements for Area 1 and Season 3, which has been identified above as the critical season. The measurements are first separated into four percentile groups (25%, 50%, 75%, and 100%) by TP, and then the data within each of the TP percentile groups is separated into four percentile groups (25%, 50%, 75%, and 100%) of TN. These segregated data are illustrated in Fig. (2). The (i, j) indices for the quantile cells $[TN]_i$ and $[TP]_j$ are shown in the background of Fig. (2), and the number of points within each quantile cell is n_{ij} .

For these measured synoptic data, correlation analysis shows that TN and TP are significantly correlated at the 90% confidence limit of the null hypothesis, which is

*

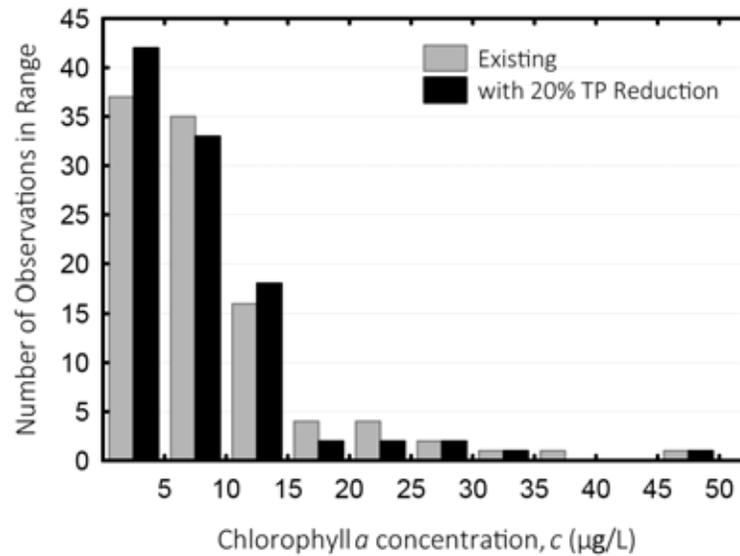


Fig. (3). Effect of TP reduction on Chla in area 1.

indicated in Column 4 of Table 3 by showing the distribution of TN as $g(TN|TP)$. Because of this correlation, TN and TP cannot be reduced independently and so specifying $(\alpha_N, \alpha_P) = (0.8, 0.8)$ would not be appropriate. Following the data analysis protocol described in the previous section, the effect of reducing TP by 20% is calculated by the following sequence of steps: take $\alpha_P = 0.8$, reduce each of the measured TP values to $\alpha_P TP$, calculate the distribution of “new measurements”, $n_{i,j}$, using Equation 7, and resample from the actual Chla measurements to determine the set of Chla values in the new measurements compared with the distribution of Chla in the actual measurements (before nutrient reduction) in Fig. (3). It is apparent from this comparison that the distribution of Chla has shifted towards lower values in response to the reduction in TP.

The data shown in Fig. (3) that is derived from resampling Chla measurements within each of the quantile cells can lead to variable c_{90} values that reflect the randomness of the Chla values observed within each (TN,TP) quantile cell. Whereas this randomness is realistic and several iterations of resampling could be done to identify the uncertainty in the estimated value of c_{90} , for determining remediation actions it is desirable to have an estimate of the expected value of c_{90} . In the present study, this expected value was determined by repeating the resampling 1000 times and then averaging the values of c_{90} ; this led to an average value with very little sensitivity to the number of resamples.

6. DISCUSSION AND CONCLUSION

High algae concentrations, as measured by Chlorophyll *a* (Chla), are commonly assumed to be symptomatic of high nutrient levels, where the influential nutrients are commonly taken as dissolved species of nitrogen and/or phosphorus, which are measured using their surrogates total nitrogen

(TN) and total phosphorus (TP). Fundamental relationships are presented showing that the probability distribution of Chla depends on two separate conditional probability distributions: (1) the probability distribution of Chla for given values of TN and TP; and (2) the probability distribution of TN for given values of TP. It is further shown that the functional forms of these distributions can be identified using the non-parametric Spearman correlation coefficients. Identification of the appropriate functional forms are essential in identifying which, if any, of the nutrients can be adjusted to influence Chla levels, and also in identifying whether TN and TP can be adjusted independently to achieve the desired outcome.

Algae control is commonly associated with nutrient control, and setting limits on Chla concentrations is usually done by setting limits on nutrients. Setting numeric nutrient limits requires that a quantitative relationship between Chla, TN, and TP be known. Conventional approaches to estimating this quantitative relationship have included both empirical and mechanistic models, with both of these approaches limited by their a-priori assumption of simplified phenomenological relationships. For example, empirical models commonly assume a linear relationship between either the raw or transformed variables and hence cannot accommodate nonlinear relationships that are not linearized by data transformation. Similarly, mechanistic models are generally based on process equations that are highly simplified. Of equal concern is that both empirical and mechanistic approaches are deterministic and so do not adequately represent the stochastic nature of the variables.

This paper presents a stochastic approach for estimating the relationship between Chla, TN, and TP that does not depend on any a-priori assumptions of the functional relationships between Chla, TN, and TP, but rather directly utilizes the measured relationship between these variables and

reflects the natural variability that is embedded in the measured data. The only fundamental assumption is that whenever TN and TP values are within a specified range of values, the corresponding Chla values are characterized by those observed when the measured TN and TP values are within the specified range. Stated in mathematical terms, a continuum of cells can be identified in the TN-TP plane such that observed Chla values are associated with each cell, and the variability in Chla values for TN and TP values within each cell can be represented by resampling the observed data. Hence, changes in nutrient concentrations can be represented simply as adjustments to measured data in the TN-TP plane, and then ascribing to these adjusted data the Chla values corresponding to their new cell locations in the TN-TP plane.

The proposed approach has been demonstrated by application to six monitoring areas within an impaired segment of the Halifax River in northeast Florida, which is eutrophic and has mean values of TN/TP in the range of 5.1–8.1. Results from these analyses show that the functional relationship between Chla, TN, and TP generally varies between monitoring area and season, with this spatial and temporal variability also occurring in the functional relationship between TN and TP. These analyses also show that using the Redfield ratio as a basis for identifying the controlling nutrient is not reliable. Although all the monitoring areas were within a single impaired segment, the results showed that relationships derived from the lumped data do not adequately represent the local relationships within the individual monitoring areas. Analysis of 90-percentile Chla values indicate that the highest values generally occur in April–September, and the most effective nutrient reduction strategies vary between monitoring areas located within the impaired segment.

The methodology described here can be used in a variety of applications, such as determining the response of Chla levels to changes in nutrient levels, determining numeric nutrient limits corresponding to specified Chla limits, determining nutrient total maximum daily loads (TMDLs), and other remediation applications.

CONFLICT OF INTEREST

The author confirms that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

Declared none.

REFERENCES

- [1] Downing J, McCauley E. The nitrogen:phosphorus relationship in lakes. *Limnol Oceanogr* 1992; 37(5): 936-45.
- [2] McCauley E, Downing J, Watson S. Sigmoid relationships between nutrients and chlorophyll among lakes. *Can J Fish Aquat Sci* 1989; 46: 1171-5.
- [3] Redfield A. On the proportions of organic derivations in sea water and their relation to the composition of plankton. In: James Johnstone Memorial Volume, ed. Daniel, R, 177-92. Liverpool, England: University Press 1934.
- [4] Seip K. Phosphorus and nitrogen limitation of algal biomass across trophic gradients. *Aquat Sci* 1994; 56(1): 16-28.
- [5] Wetzel R. *Limnology: Lake and River Ecosystems*. New York: Academic Press 2001.
- [6] Håkanson L, Boulion V. *The Lake Food Web—Modelling Predation and Abiotic/Biotic Interactions*. The Netherlands: Backhuya Publishers 2002.
- [7] Guildford S, Hecky R. Total nitrogen, total phosphorus, and nutrient limitation in lakes and oceans: Is there a common relationship?. *Limnol Oceanogr* 2000; 45:1213-23.
- [8] Sverdrup H. On conditions for the vernal blooming of phytoplankton. *J du Conseil Int pour l'Exploration de la Mer* 1953; 18: 287-95.
- [9] Townsend D, Keller M, Sieracki M, Ackleson SG. Spring phytoplankton blooms in the absence of vertical water column stratification. *Nature* 1992; 350: 59-62.
- [10] Håkanson L, Eklund J. Relationships between Chlorophyll, Salinity, Phosphorus, and Nitrogen in Lakes and Marine Areas. *J Coastal Res* 2010; 26(3): 412-23.
- [11] Håkanson L. Sediment sampling in different aquatic environments: statistical aspects. *Water Resour Res* 1984; 20: 41-6.
- [12] Press W, Flannery B, Teukolsky S, Vetterling W. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, United Kingdom: Cambridge University Press 1992.
- [13] Magley, W. Draft TMDL Report, Nutrient TMDL for Halifax River, WBID 2363B. – EPA-600/S1-81-006. Tallahassee, Florida : Florida Department of Environmental Protection 2013.
- [14] Chin D. *Water-Quality Engineering in Natural Systems*. 2nd ed. Hoboken, New Jersey: John Wiley & Sons 2013.

Received: March 25, 2015

Revised: April 23, 2015

Accepted: April 23, 2015

© David A. Chin; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.