WATER RESOURCES BULLETIN
VOL. 20, NO. 3 AMERICAN WATER RESOURCES ASSOCIATION JUNE 1984

SOURCES OF VARIABILITY IN PHOSPHORUS AND CHLOROPHYLL AND THEIR EFFECTS ON USE OF LAKE SURVEY DATA[1]

Matthew F. Knowlton, Mark V. Hoyer, and John R. Jones[2]

ABSTRACT: Summer lake survey measurements of total phosphorus (TP) and chlorophyll *a* (CHLa) from 188 reservoirs and natural lakes in the midwest were analyzed to determine the magnitude of major sources of variability. Median variance among replicate samples collected at the same location and time was about 7-8 percent of the mean for both TP and CHLa. Median observed temporal variability within summers was 27 percent of the mean for TP and 45 percent of the mean for CHLa. Median values of year-to-year variance in average TP and CHLa were 22 percent and 31 percent of the mean, respectively. A range of approximately two orders of magnitude was observed among individual estimates of variance in each of these categories. The magnitude of observed temporal variability was affected only slightly by variance among replicate samples on individual days and was weakly correlated with the length of time during which samples were collected from individual lakes. Observed temporal variation was similar between reservoirs and natural lakes when variances were calculated with log-transformed data. The magnitude of temporal and year-to-year variance can severely limit the power of statistical comparisons of TP and CHLa means, but has less effect on establishing relative rankings of lake means. Sources and relative magnitude of variability are important in the use of TP and CHLa data in regression models and in the planning of lake surveys and subsequent data analysis.
(KEY TERMS: total phosphorus; chlorophyll *a*; sampling error; temporal variation; trophic state; regression analysis.)

INTRODUCTION

Mean values of total phosphorus (TP) and chlorophyll *a* (CHLa) are widely used for modeling large-scale limnological relations (e.g., Sakamoto, 1966; Vollenweider, 1975; Jones and Bachmann, 1976; Smith, 1979; Oglesby, 1977) and for the assessment of water quality in lakes (e.g., Likens, 1975; Edmondson and Lehman, 1982). Use of TP and CHLa means for such purposes is affected by the varying degree of statistical certainty in their measurement. Different sources of analytical error, spatial and temporal patchiness in phytoplankton and nutrient distribution, sampling protocols, and the duration and intensity of sampling in individual studies may have combined effects on the precision of TP and CHLa averages (Reckhow, 1979; Walker, 1979; Stauffer, *et al.,* 1979; Prepas and Rigler, 1982). Identifying quantitatively significant sources of uncertainty in the measurement of TP and CHLa means is important in the design of limnological and water quality studies (Reckhow, 1979; Thornton, *et al.,* 1982; Stauffer, 1981) and interpretation of data from large-scale limnological investigations (Collins and Sprules, 1983). Increasing use of TP and CHLa data for purposes of water quality assessment and allocation of lake restoration efforts is a major impetus for examining, more closely, the degree to which estimated means of TP and CHLa can be treated as deterministic values to characterize conditions in a given water body.

Much of the data presently used in large-scale modeling and in general assessments of water quality are summer means of TP and CHLa obtained in multi-lake surveys designed to achieve the broadest possible scope in terms of the numbers, geographical range and limnological variety of lakes sampled (e.g., Moyle, 1956; Jones and Bachmann, 1978a; Lillie and Mason, 1983). In keeping with the scale of these studies, sampling in individual lakes has often been limited to collection of surface water samples at some representative location on the lake several times in late spring and summer of one or a few years. Such data represent a body of information whose appropriate use and interpretation may depend heavily on variability characteristics, but about which no general quantitative assessment of variability has yet been made.

This paper is a preliminary assessment of the magnitude of variability affecting measurement and use of summer TP and CHLa data from lake surveys. Included are analyses of variability in TP and CHLa measurements in surface water samples on a given day and over time and an assessment of the use of lake survey data for comparing and ranking means of TP and CHLa. Some effects of variability on TP-CHLa regression models and planning and execution of lake surveys are also discussed.

DATA BASE

Data from 67 natural lakes and 121 reservoirs located in Missouri, Iowa, and Minnesota, representing a range of trophic conditions, were used in the analysis (Table 1). Data are from

[1]Paper No. 84110 of the *Water Resources Bulletin.* (Contribution of the Missouri Agricultural Experiment Station Journal Series No. 8954).
[2]School of Forestry, Fisheries and Wildlife, 112 Stephens Hall, University of Missouri, Columbia, Missouri 65211. (Present address for Hoyer: Center for Aquatic Weeds, University of Florida, Gainesville, Florida 32611.)

surface samples collected at the same location on each water body on 3 to 12 dates between 1 June and 15 September. Individual water bodies were sampled for 1 to 6 summers (during 1971 to 1981), yielding a total of 313 data sets (a data set is considered as all data from one summer at one location). Most of the Missouri and Iowa data sets (219 of 245) were based on replicate samples (usually 3) collected separately within a 1000 m^2 area on each date. The remainder of the data sets and the data from Minnesota lakes (68 data sets) were based on single daily measurements of TP and CHLa. Natural lakes were sampled at the deepest part of the lake; reservoirs were sampled in the lacustrine zone. No attempt was made to characterize the longitudinal gradations frequently observed in large reservoirs (Thornton, *et al.*, 1982) or to assess the adequacy of surface samples from a single location in estimating surface water TP and CHLa for the entire water body (Stauffer, 1981).

TABLE 1. Mean, Standard Deviation (SD), and Range of TP and CHLa (mg/m^3) for 121 Reservoirs and 67 Natural Lakes (Shapiro and Pfannkuck, 1974; Jones and Bachmann, 1978; Noonan, 1979; Osgood, 1981, 1982; Hoyer and Jones, 1983). Six stations on Lake of the Ozarks, Missouri, differ in limnological characteristics and were treated as separate water bodies (Jones and Novak, 1981).

	Mean	SD	Range
TP			
Reservoirs	46	37	6-254
Natural Lakes	93	90	20-444
All Data	63	65	6-444
CHLa			
Reservoirs	23	25	1-167
Natural Lakes	51	43	2-185
All Data	33	25	1-185

A random effects analysis of variance procedure (Snedecor and Cochran, 1980: p. 238) was used on the data sets from each lake to estimate the magnitude of variance of TP and CHLa from three sources: variance among replicate samples on a given day (sampling error – σ_s^2), variance among true daily values of TP or CHLa in a given summer (temporal variance – σ_d^2), and variance among true summer means for a given lake (year-to-year variance – σ_y^2). Application of this analysis of variance model requires the assumption that the distributions of summer means around the true lake means (a lake mean in this context is considered to be the true mean of summer TP or CHLa for the period of years about which inferences are to be made), daily values around the summer mean, and individual measurements around the true daily values of TP and CHLa are normally and independently distributed. The normality assumption is often not correct with data of this sort (Walmsby, 1984) and the time series nature of these data may result in nonindependence of the observations (Reckhow and Chapra, 1983). Given the scope and heterogeneity of the present data we cannot assess the degree to which failure of these assumptions effects the estimation and use of these variance components. While it is likely that these effects will introduce biases in the estimation of variance components for individual data sets we will assume that the extent of these effects on the general patterns of variability in these data are small.

Estimates of true variance were calculated from squares of standard deviations of replicate TP or CHLa on individual days (S_s^2), daily TP or CHLa in a given summer (S_d^2), and summer means of TP or CHLa from individual lakes (S_y^2) after the equations in Table 2. Estimations from these equations may be biased when applied to data which are unbalanced with respect to numbers of replicates collected on a given day (n) or numbers of sampling dates per year (d). For the present analysis, however, calculation of variance components with techniques which account for lack of balance present in some of these data (e.g., Snedecor and Cochran, 1980: p. 246) usually resulted in only minor changes in the estimates of variance components and had almost no effect on median values and ranges of these estimates from all data sets. Because our intent here is to show general patterns in sources of variation in TP and CHLa data rather than to make rigorous statistical tests, we have used this approximate technique for the sake of simplifying calculations and relating components of variation to readily recognized measures of observed variability (squares of standard deviations).

About one third of the present data are based on single, rather than replicate, daily measurements of TP and CHLa and provided no measurement of observed sampling error (S_s^2). For these lakes, estimates of true temporal variance (σ_d^2), and true year-to-year variance (σ_y^2) were based on the median values of observed sampling error (S_s^2) from the remaining data sets. In the present data observed variation at each level was correlated with the corresponding estimates of daily, summer, or lake means (Figure 1). To satisfy the constancy of variance assumption of parametric statistical analysis (Snedecor and Cochran, 1980: p. 291), the correlation between variances and means was removed by transforming raw data to base 10 logarithms prior to analysis. Because of the unfamiliar scale of variance estimates expressed in logarithmic units, the magnitude of variances are described in the text by the statistic:

$$\% \text{ variance} = 50[10^{\sqrt{a}} - 1/10^{\sqrt{a}}] \quad (1)$$

where a is a value of S_s^2, S_d^2, S_y^2, $\hat{\sigma}_s^2$, $\hat{\sigma}_d^2$, or $\hat{\sigma}_y^2$. These values express variance as a percent of the mean and are approximately equal to coefficients of variation calculated from untransformed data (Snedecor and Cochran, 1980: p. 37). Values of % variance are included parenthetically with logarithmic estimates of variance in the text. Median values of $\hat{\sigma}_s^2$, $\hat{\sigma}_d^2$, and $\hat{\sigma}_y^2$ were used to estimate the effect of varying the numbers of replicate samples of TP or CHLa on a given date (n), the number of sampling dates per summer (d), and the number of

TABLE 2. Estimators of Observed Variance and Components of True Variance for Individual Lakes.

Source	Observed Variance	Variance Components	Variance Component Estimator
Sampling Error	$S_s^2 = \frac{\Sigma(X - \bar{X})^2}{n - 1}$	σ_s^2	$\hat{\sigma}_s^2 = \frac{\Sigma\Sigma S_s^2}{\Sigma d}$
Temporal Variance	$S_d^2 = \frac{\Sigma(\bar{X} - \bar{\bar{X}})^2}{d - 1}$	$\sigma_d^2 + \frac{\sigma_s^2}{n}$	$\hat{\sigma}_d^2 = \frac{\Sigma S_d^2}{y} - \frac{\hat{\sigma}_s^2}{\bar{n}}$
Year-to-Year Variance	$S_y^2 = \frac{\Sigma(\bar{\bar{X}} - \bar{\bar{\bar{X}}})^2}{y - 1}$	$\sigma_y^2 + \frac{\sigma_d^2}{d} + \frac{\sigma_s^2}{dn}$	$\hat{\sigma}_y^2 = S_y^2 - \frac{\hat{\sigma}_d^2}{\bar{d}} - \frac{\hat{\sigma}_s^2}{\bar{d}\bar{n}}$

X = Log_{10} transformed TP or CHLa of a single sample. $\bar{X}$ = mean of X for a given day (daily value).
$\bar{\bar{X}}$ = Mean of $\bar{X}$ for a given summer (summer mean).
$\bar{\bar{\bar{X}}}$ = Mean of $\bar{\bar{X}}$ all summers (lake mean).
$\bar{n}$ = Mean number of replicates for all sample dates.
$\bar{d}$ = Mean number of sample dates per summer.

years sampled (y) on the power of comparisons between pairs of summer or lake means. Estimations of power were calculated according to Snedecor and Cochran (1980: p. 103) for t-tests of two independent samples. Other statistical analyses (noted in the text) were performed using programs of the Statistical Analysis System (1982) or by procedures referenced in the text. Unless otherwise stated all tests were conducted at $\alpha = 0.05$.

RESULTS AND DISCUSSION

Variability of TP and CHLa Measurements

Summer TP and CHLa data from lake surveys provide information about three general sources of variation which are important to the interpretation of these data and the design of limnological sampling programs. They are: 1) sampling error – variation in replicate measurements of TP and CHLa from the same time and general location that results from small-scale patchiness in the surface waters and analytical imprecision; 2) temporal variability – day-to-day fluctuations in surface water concentrations of TP and CHLa in a particular summer at the location sampled; and 3) year-to-year variation – variation of true summer means of TP or CHLa around the lake mean for years of interest. Sampling error and temporal variation determine the amount of data (i.e., number of samples per day, n, and number of days per summer, d) necessary to measure the true summer mean of TP and CHLa with a given precision. Year-to-year variation greatly affects the suitability of data from one or several years for describing the long-term condition of a lake – a goal implicit in the collection of data for evaluating lake trophic status.

Variation among estimated means of TP and CHLa is affected by the magnitude of variation in the true means being estimated and by the precision of these estimates. For example, variance observed among estimated daily values of TP or CHLa from a given lake in a particular summer (S_d^2) depends not only on true variance among daily TP or CHLa values (σ_d^2) but also on sampling error (σ_s^2) and numbers of replicate TP or CHLa measurements made on each sample date (n). Likewise, precision of estimated summer means, which is itself dependent on temporal variance (σ_d^2), sampling error (σ_s^2), number of sample dates per summer (d), and number of replicate samples per sample date (n), may greatly influence observed year-to-year variability (S_y^2).

In these data sets median observed temporal variation (S_d^2) for TP and Chla were 0.0134 (27 percent) and 0.0375 (45 percent), respectively (Table 3). These variance estimates correspond to an approximate median range (maximum value/minimum value) of 1.8 fold for TP and 2.5 fold for CHLa. Infrequent sampling, which is characteristic of most of these data sets (mean = 4 samples per summer) is unlikely to detect the highest and lowest values that actually occurred. For example, the median ratio of maximum to mean CHLa in these data set was about 1.5. Jones, *et al.* (1979), found the average of this ratio to be about 1.7 for lakes sampled weekly or biweekly.

Median estimates of true temporal variance ($\hat{\sigma}_d^2$) for these data sets were 0.0137 (27 percent) and 0.0356 (45 percent), respectively, for TP and CHLa (Table 3). The close agreement between median observed variance (S_d^2) and median estimates of true variance ($\hat{\sigma}_d^2$) indicates that precision of daily estimates of TP and CHLa generally had little influence on observed temporal variation. This similarity was not surprising because median values of observed sampling error (S_s^2) were only 0.0010 (7 percent) for TP and 0.0012 (8 percent) for CHLa. Observed year-to-year variation, however, was usually much greater than estimates of true year-to-year variance. Median values of observed year-to-year variation (S_y^2) were 0.0089

(22 percent) for TP and 0.0180 (31 percent) for CHLa, while median values of true year-to-year variation ($\hat{\sigma}_y^2$) were 0.0032 (13 percent) and 0.0072 (20 percent), respectively (Table 3). This finding indicates that much of the variation among observed summer means from a given lake could result from imprecision in these estimates rather than year-to-year fluctuations in true summer averages of TP and CHLa.

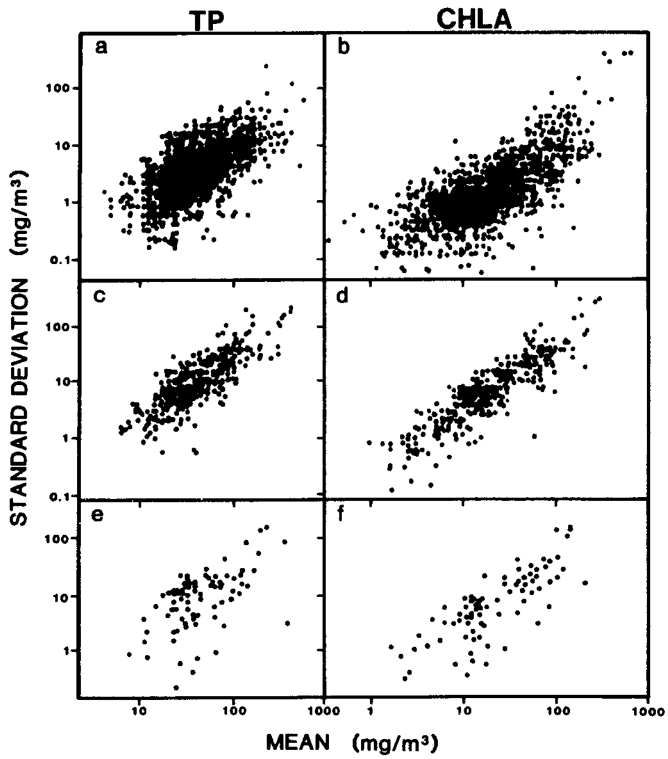

Figure 1. Relations Between Means and Standard Deviations (SD) of TP and CHLa Calculated with Untransformed Data. A and B – daily means and SD of replicate measurements; C and D – summer means and SD of daily values of TP and CHLa for individual data sets; E and F – lake means and SD of summer means from individual lakes.

Observed values of variance components were extremely heterogeneous within and among lakes covering a 2-3 order of magnitude range (Figure 2). Within lakes individual measurements of observed sampling error (S_s^2) and temporal variance (S_d^2) varied widely among days and years. For example, although estimates of true sampling error ($\hat{\sigma}_s^2$) exceeded true temporal variance ($\hat{\sigma}_d^2$) for less than 15 percent of lakes with replicated daily TP and CHLa measurements, data sets from over half of these lakes (85 of 148) contained at least one individual measurement of sampling error (S_s^2) which exceeded observed temporal variance (S_d^2) for TP, CHLa, or both.

One source of heterogeneity of variance estimates in these data is the low statistical precision with which variance components are estimated from relatively small numbers of measurements (Snedecor and Cochran, 1980: p. 74, 245). Imprecision in estimation of true sampling error (σ_s^2) and temporal variance ($\hat{\sigma}_d^2$), for example, resulted in our obtaining negative values for about 25 percent of our estimates of true year-to-year variance ($\hat{\sigma}_y^2$, Table 3). It is likely, however, that much of the heterogeneity of variance components represented in these data results from diversity in the limnological conditions that cause variability in TP and CHLa. Because of the uncertainty in estimating components of variance for individual lakes and the variety of conditions that affect variation of TP and CHLa in lakes, we have restricted our present analyses to large-scale patterns represented by average or median variance of this large and diverse group of water bodies.

TABLE 3. Medians and Ranges of Observed Sampling Error (S_s^2), Observed Temporal Variance (S_d^2), and Observed Year-to-Year Variance (S_y^2) and Estimates of Corresponding Components of True Variance ($\hat{\sigma}_s^2, \hat{\sigma}_d^2, \hat{\sigma}_y^2$) for TP and CHLa. Parenthetic values are corresponding estimates of % variance.

	TP		CHLa	
	Median	Range	Median	Range
S_s^2	0.0010 (7)	0.0000-0.1270 (6-92)	0.0012 (8)	0.0000-0.3414 (0-179)
$\hat{\sigma}_s^2$	0.0019 (10)	0.0001-0.0438 (2-50)	0.0027 (12)	0.0001-0.1215 (2-89)
S_d^2	0.0134 (27)	0.0000-0.2182 (0-130)	0.0357 (45)	0.0001-0.7073 (2-340)
$\hat{\sigma}_d^2$	0.0137 (27)	−0.0017-0.2173 (0-129)	0.0356 (45)	−0.0087-0.4693 (0-232)
S_y^2	0.0089 (22)	0.0000-0.1016 (2-80)	0.0180 (31)	0.0000-0.8282 (0-400)
$\hat{\sigma}_y^2$	0.0032 (13)	−0.0365-0.0993 (0-79)	0.0072 (20)	−0.0706-0.8183 (0-395)

An easily recognized pattern of variation in summer TP and CHLa measurements is their relation to mean TP and CHLa when both means and variance are calculated from untransformed data. Standard deviations, expressed as mg/m^3, are correlated with corresponding means of replicates, summer means, and lake means, with scatter in this relation also increasing with the mean (thus the logarithmic scales in Figure 1). In a study of year-around variability of CHLa in South African impoundments, Walmsley (1984) found a similar relation of variance to the mean.

Calculation of variance components from log-transformed data expresses variation as a proportion of the mean. This approach removes the correlation of variance to the mean and reveals different patterns of variability. For example, TP values in lakes are generally about twice as great as CHLa

(Table 1) and a similar difference occurs between the variation observed in these parameters when variation is expressed in mg/m^3 (Figure 1). When variation is calculated with log-transformed data and expressed as a percent of mean, however, temporal variance and year-to-year variance of TP is about one-third less than CHLa and sampling error for the two parameters is nearly equal (Table 3). In view of the large number of factors that can influence algal biomass in lakes, the greater average % variability of CHLa over time is not surprising.

Another factor affecting observed temporal variation is the duration of sampling in individual studies. The time span during which TP and CHLa data were collected in individual surveys ranged from 14 to 111 days (mean = 55 days) and was weakly correlated with the amount of variation observed among daily TP and CHLa values (S_d^2). On the average, observed temporal variance (S_d^2) increased by 0.0024 (11 percent) with each week the sampling period was extended (Figure 3). Differences in sampling period, however, accounted for only 1 percent (CHLa) to 10 percent (TP) of the variation in S_d^2 values. This uncertain effect of length of sampling period suggests that temporal variability of TP and CHLa in these data sets usually results from fluctuations of these parameters over a few weeks as opposed to more gradual seasonal trends like the cessation of spring algal blooms or the summer decline of TP observed in some water bodies (Jones and Bachmann, 1978b; Lillie and Mason, 1983).

Both natural lakes and reservoirs were included in these analyses. On the average, temporal variance estimates from untransformed data (standard deviations) were about twice as great for natural lakes as for reservoirs because of the higher average TP and CHLa in these water bodies (Table 1) and the relation of variance to the mean (Figure 1). When variance was determined from log-transformed data the difference between the two lake-types was greatly reduced (Table 4). Although average observed temporal variability (S_d^2) for natural lakes somewhat exceeded that of reservoirs, the extensive overlap of these data suggests that variability of TP and CHLa in summer does not reflect the general morphological and hydrologic differences between lakes and reservoirs (Thornton, *et al.*, 1982). This overlap also suggests that differences in temporal variability of TP and CHLa are not predictable from classifications of these lake-types when variance is stabilized for differences in means.

Comparing and Ranking TP and CHLa Means

Summer means and lake means of TP and CHLa from lake surveys are typically used to document the present conditions of individual lakes for the purpose of monitoring changes over time and for making comparisons among groups of lakes. Comparisons made with these data can be grouped into two general types which we will call statistical comparisons and rankings. Statistical comparisons involve submitting the means to statistical tests like t-tests of pairs of means or analysis of variance procedures (e.g., Duncan's multiple range test) for comparing several means simultaneously. Rankings make no rigorous inferences of the magnitude of differences among means but attempt only to arrange data from different lakes in hierarchical order (i.e., rank the trophic status of lake within a given region).

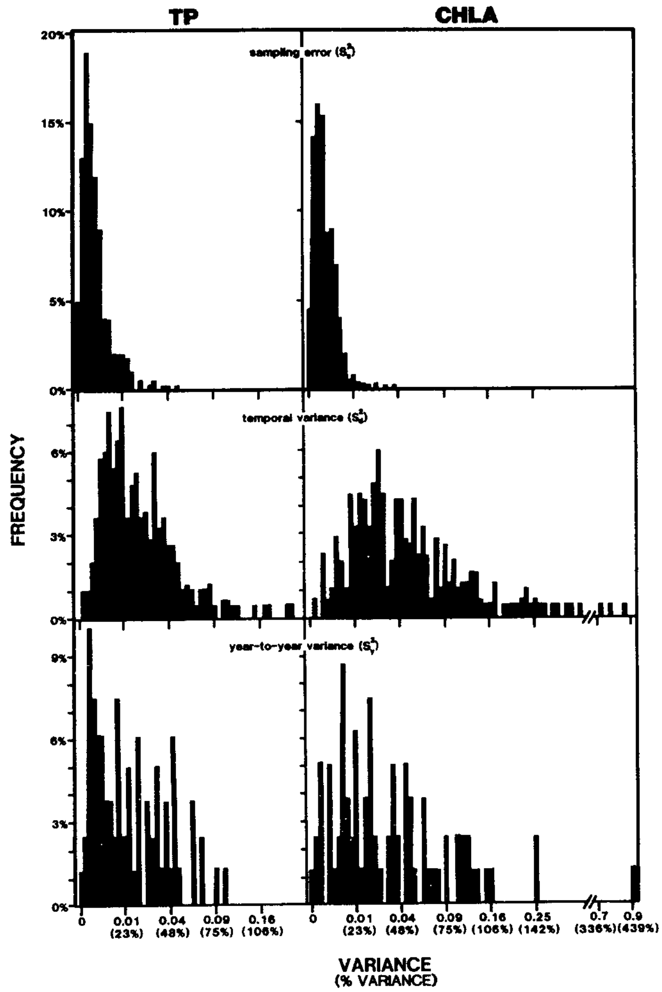

Figure 2. Frequency Distribution (%) of 887 Values of Observed Sampling Error (S_s^2), 313 Values of Observed Temporal Variance (S_d^2), and 81 Values of Observed Year-to-Year Variance (S_y^2) of TP and CHLa Calculated From These Data Sets. Estimates of % variation are in parentheses. Values of sampling error comprising less than 0.3 percent of the total are not shown.

Summer and lake means from lake surveys, such as those included in this analysis, are not always amenable to rigorous statistical comparison because the variances affecting these measurements are extremely heterogeneous. Logarithmic transformation stabilizes variance for differences in means, but does not reduce heterogeneity. Also, it is unlikely that the assumption of independence of observations is met in these data, although it is possible that fluctuations in TP and CHLa were sufficiently rapid with respect to the intervals between sampling that time series dependence in these data may be a minor effect. Nonhomogeneous variance and nonindependence of observations reduce the reliability of parametric statistical tests and necessitate use of nonparametric analyses. Also, much of the data available from lake surveys represents lakes in only one year, as did data for over half of the lakes in this analysis. Such data provide no estimate of year-to-year variance (σ_y^2) and no statistical degrees of freedom for testing differences in lake means.

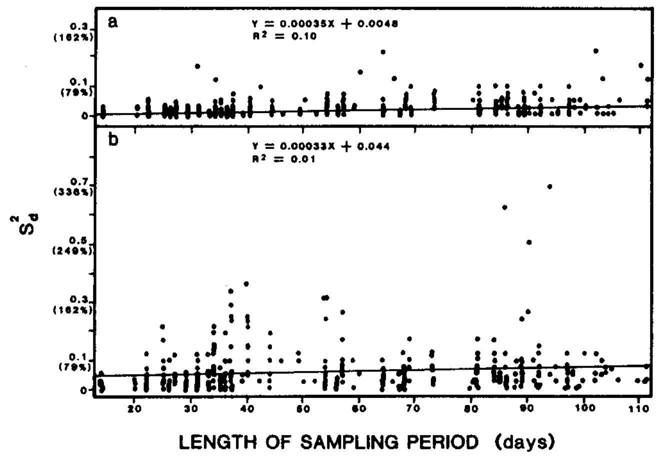

Figure 3. Relations of Observed Temporal Variance (S_d^2) to Length of Sampling Period (number of days between first and final sampling in a given summer) for TP (a) and CHLa (b). Significance levels of the fitted regression were $\alpha = 0.0001$ (a) and $\alpha = 0.0545$ (b). Both logarithmic variance estimates and % variance values (in parentheses) are displayed on the vertical axis.

Another deficiency of statistical comparisons with these data is their relatively low power. Power of a statistical test is the probability that it will detect a significant difference among estimates of means whose true values are different (Snedecor and Cochran, 1980: p. 68). Power is a function of the statistical test, its significance level (usually held at 5 percent), size of differences between true values of compared means, the number of means compared and the precision of estimated means used in comparisons. This precision is a function of the magnitude of variances affecting the estimation (e.g., σ_s^2, σ_d^2, and σ_y^2) and the amount of data (e.g., n, d, and y) from which estimates were made.

As a simple example of how these characteristics affect comparisons of summer means and lake means of TP and CHLa we used median values of $\hat{\sigma}_s^2$, $\hat{\sigma}_d^2$, and $\hat{\sigma}_y^2$ from our data to calculate the power of a comparison of two means at different levels of n, d, and y with a simple t-test (Snedecor and Cochran, 1980: p. 103-104). We chose, for the example, to calculate power for comparisons of summer means whose true values differed by a factor of 1.5 fold (50 percent) and summer and lake means whose true values differed by two-fold (100 percent) because differences of these magnitudes are sometimes used to distinguish relative trophic status (e.g., Carlson, 1977). We found that increasing numbers of replicate measurement of TP or CHLa (n) on a given day had little effect on power of comparison of summer means or lake means. For example, increasing n from 1 to 10 samples per day would increase power of these hypothetical tests of summer means and lake means by less than 3 percent, so we have held n constant at one sample per day in these analyses.

TABLE 4. Comparison of Observed Temporal Variation of Total Phosphorus and Chlorophyll *a* Between Reservoirs and Natural Lakes with Data Transformed to Base 10 Logarithms (S_d^2 values) and Untransformed Data (standard deviations in mg/m^3). Parenthetic values are estimates of % variance corresponding to S_d^2 values or coefficients of variation (standard deviation X 100/mean) corresponding to standard deviations.

Lake Type	TP			CHLa		
	Mean	Median	Range	Mean	Median	Range
			LOGARITHMIC			
Reservoirs	0.022*	0.011	0.000-0.218	0.059	0.031	0.000-0.364
	(35)	(24)	(2-129)	(59)	(42)	(2-188)
Natural Lakes	0.027*	0.018	0.000-0.212	0.073	0.039	0.002-0.707
	(39)	(31)	(1-127)	(66)	(47)	(10-339)
			UNTRANSFORMED			
Reservoirs	14.3*	7.9	0.5-213.4	9.8*	5.2	0.1-163.5
	(29)	(23)	(2-143)	(44)	(39)	(2-147)
Natural Lakes	27.2*	16.0	0.6-193.5	16.5*	13.7	0.5-278.4
	(32)	(29)	(1-105)	(47)	(44)	(10-145)

*Significant difference between reservoirs and natural lakes using Wilcoxin rank sum test (S.A.S., 1982).

From this analysis we found that the average amount of data represented in these data sets (four samples per summer in two years for each lake) would provide only a 37 percent probability of detecting a two-fold difference in CHLa (Figure 4). Thus data like those analyzed here may often be too imprecise for application of numerical indices that rely on distinguishing two-fold differences among lakes as a criterion for evaluating relative trophic status. To obtain a 90 percent probability of detecting a two-fold difference between two lake means of CHLa might require three samples per year for five years or seven samples per year for four years. The variability of TP as a percent of the mean is less than CHLa, so differences among lake means of TP are more readily distinguished. A sampling frequency of four samples per summer for two years would provide a 73 percent probability of detecting a two-fold difference in TP lake means (Figure 4).

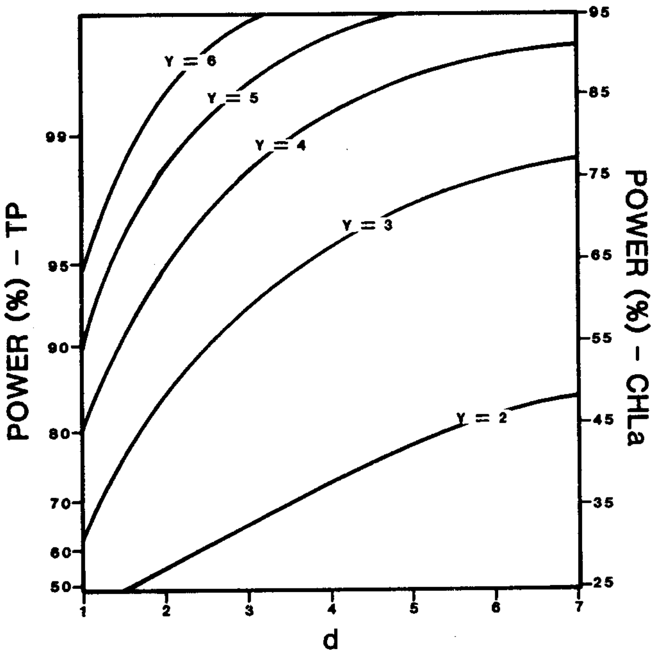

Figure 4. Approximate Power of a t-Test (α = 0.05) of Two Lake Means of TP or CHLa as a Function of d, the Number of Sample Dates Per Summer, and y, the Number of Summers Sampled Per Lake. Power was calculated from: Power (%) = 100(0.5 + probability of Z), where Z (the standard normal deviate from Table A3 of Snedecor and Cochran, 1980: p. 468) = $\{[a-1]^{1/2}[\log_{10}(b)]/(1.41c)\} - 1.96$. Here a = number years samples per lake (y), b = difference (as a multiplication factor) between the true lake means whose estimated values are being compared (two fold in this example), and $c = [\hat{\sigma}_y^2 + \hat{\sigma}_d^2/d + \hat{\sigma}_s^2/nd]^{1/2}$. The number of replicate samples per day, n, is held constant at one and it is assumed that $\hat{\sigma}_y^2$, $\hat{\sigma}_d^2$, and $\hat{\sigma}_s^2$ equal 0.0032, 0.0137, and 0.0010, respectively, for TP; and 0.0072, 0.0375, and 0.0012, respectively, for CHLa. These estimates are derived from an approximate relation between numbers of samples and power for t-tests of pairs of independent means (Snedecor and Cochran, 1980: p. 104).

Comparisons of means from individual summers are somewhat more powerful than comparisons of lake means. A sampling frequency of four sample dates per summer would provide a power of 86 percent and 50 percent, respectively, for detecting two-fold differences among pairs of TP summer means and CHLa summer means, but only a 45 percent and 20 percent power of detecting 1.5 fold differences (Figure 5). These examples pertain to single comparisons of two means and may greatly overestimate power of simultaneous statistical comparisons of large numbers of means produced by a typical lake survey. Also, because they are based on median variances, we can assume that examples shown here would overestimate power for about half of the comparisons that might be made with these data.

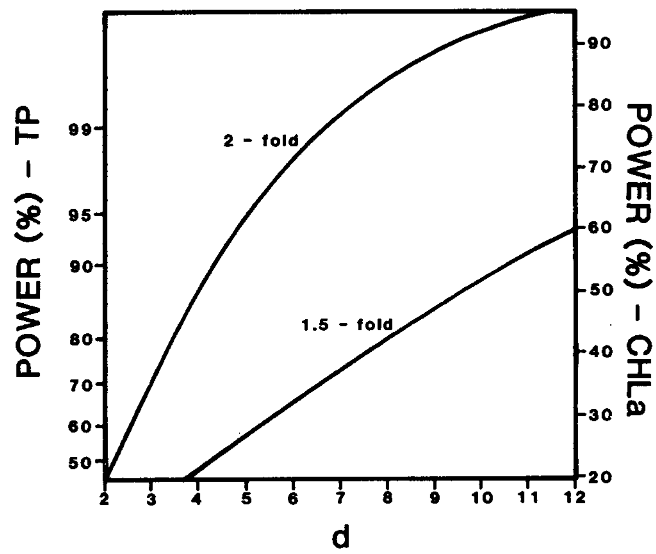

Figure 5. Approximate Power of a t-Test (α = 0.05) Comparing Summer Means of TP or CHLa as a Function of d, the Number of Sampling Dates Per Summer. Calculations of power are the same as for Figure 4 except that a = number of days sampled per summer (d), b is 2 or 1.5, and $c = [\hat{\sigma}_d^2 + \hat{\sigma}_s^2/n]^{1/2}$.

A far less demanding use for lake survey data is establishing rankings of lake means of TP and CHLa. We can consider that such rankings are successful if estimated means are misranked relative to rankings of their true values only if differences among true lake means are comparatively small. We evaluated the ability of lake survey data to correctly rank a given pair of lake means of TP and CHLa by calculating the probability that estimated means, based on different amounts of data (i.e., a range of values of y and d holding n constant at one sample per day), would correctly rank true lake means of TP and CHLa that differed by 25 percent, 50 percent, and two-fold (Table 4). Assuming values of sampling error, temporal variance, and year-to-year variance equal to or less than the medians for these data sets, a single sample of TP or CHLa from one year would provide an 84-94 percent probability of correctly ranking two lake means whose true values differed by two-fold, but only a 63-69 percent probability of correctly ranking lakes that differed by 25 percent (Table 5). With data from four days in each of two years (the averages of d and y for these data) estimated means would be almost certain to correctly rank the average TP and CHLa values of lakes differing by two-fold and would provide a 77-88 percent probability of correctly ranking lakes differing by 25 percent. As was true with power of statistical tests, the probability of correctly ranking any pair of lakes will decline as the number of lakes being ranked increases, but this analysis suggests that even lake survey data based on few samples are unlikely to misrank lakes that differ by two-fold or more.

TABLE 5. Probability of Correctly Ranking Pairs of Lake Means of TP and CHLa Whose True Values Differ 1.25 Fold, 1.5 Fold, and 2 Fold as a Function of Numbers of Samples Per Year (d) and Number of Years Sampled (y). Probabilities are calculated by the relation: Probability (%) of correct ranking = 100(0.5 + probability of Z), where Z is the normal standard deviate from Table A3 of Snedecor and Cochran (1980: p. 468) and is determined as $Z = D(2S/y)^{1/2}$. In this relation D is the difference as a multiplication factor between the true means and $S = \hat{\sigma}_y^2 + \hat{\sigma}_d^2/d + \hat{\sigma}_s^2/dn$ where values of $\hat{\sigma}_y^2$, $\hat{\sigma}_d^2$, $\hat{\sigma}_s^2$ are medians for these data and n is held at one sample per day.

		D = 1.25		D = 1.5		D = 2.0	
y	d	TP	CHLa	TP	CHLa	TP	CHLa
1	1	69	63	82	72	94	84
1	4	79	70	93	88	99	95
1	8	83	74	96	87	99+	97
2	1	76	67	90	80	99	92
2	4	87	81	98	95	99+	99+

Regression Analyses with TP and CHLa Data

An important use of TP and CHLa data from lake surveys and other limnological studies has been the construction of regression models for the prediction of phytoplankton abundance. These models have provided insight into the expression of nutrient limitation in lakes and are among the only tools currently available for making large-scale predictions about phytoplankton abundance (Klaff and Knoechel, 1978). Nonetheless, because of a variety of additional factors that can influence the relation of CHLa to TP (e.g., CHLa content of algal cells, availability of other nutrients, zooplankton grazing, light limitation, temperature effects, phosphorus availability, and others), TP-CHLa regression models have shown a limited ability to predict average phytoplankton abundance (e.g., Canfield and Bachmann, 1981) and their ability to predict phytoplankton abundance at any given time is weaker still.

Daily measurement of TP and CHLa in summer often fluctuate widely around their mean values – the average range of daily values of CHLa in the present data sets, for example, was about three-fold and in individual data sets often exceeded ten-fold. Thus, it is important in development and assessment of TP-CHLa models to distinguish between variation affecting mean values and variation affecting values of TP and CHLa on individual days. Ignoring the difference between these levels of variability can lead to erroneous interpretation of the relation these models depict. For example, in a recent paper, Smith and Shapiro (1981) interpret data collected by Megard (1972) from 11 locations on Lake Minnetonka as showing effects of seasonality on relations of CHLa to TP. Data in Figure 2a and 2b from their paper (represented here by solid circles in Figure 6a and 6b) show a strong relation between means of TP and CHLa in summer (Figure 6b), but a weak and scattered relation for unaveraged data which included measurements from spring, summer, and fall (Figure 6a). The authors suggest that the difference in scatter between the two plots is a seasonal effect and state "It is clear . . . [that the strong relation between TP and CHLa in summer] . . is obscured in Figure [6a] by the inclusion of spring and fall data." The dramatic difference in scatter between Figures 5a and 5b, however, is mostly an effect of averaging. Unaveraged data from summer alone (Figure 6c) show a degree of scatter similar to that in data from spring, summer, and fall combined (Figure 6a), while means from spring, summer, and fall (open circles in Figure 6b) are as strongly correlated as means from summer alone.

This discussion does not imply that seasonality has no influence on TP-CHLa relations in Lake Minnetonka; indeed, ratios of CHLa to TP in these data are significantly correlated with water temperature (CHLa/TP = 0.025 C^o – 0.038, r = 0.66). This example does, however, illustrate the degree of uncertainty involved in using relations among means to make specific inferences about conditions that vary through time and depicts the importance of considering the magnitude of this variation in development of models. The CHLa-TP regressions developed from the present data sets (Table 6) suggest that factors resulting in variation in limnological conditions through time may also influence the relation between phytoplankton and nutrients. In these data relations between summer means of TP and CHLa differ among data sets with different levels of observed temporal variation. Regressions fitted for data sets with S_d^2 values for TP or CHLa below the median values for these data differed significantly (F-test – Neter and Wasserman, 1974: p. 160), with greater slopes and considerably higher coefficients of determination, from models derived from data sets with variability above the median (Table 6). It is not presently clear whether this difference is a biological effect which relates variability of limnological conditions to the expression of nutrient limitation or an effect of heterogeneous variances in data on estimation of regression parameters. In either case this result points out the need for considering the variance characteristics of limnological data in the development and interpretation of large-scale models.

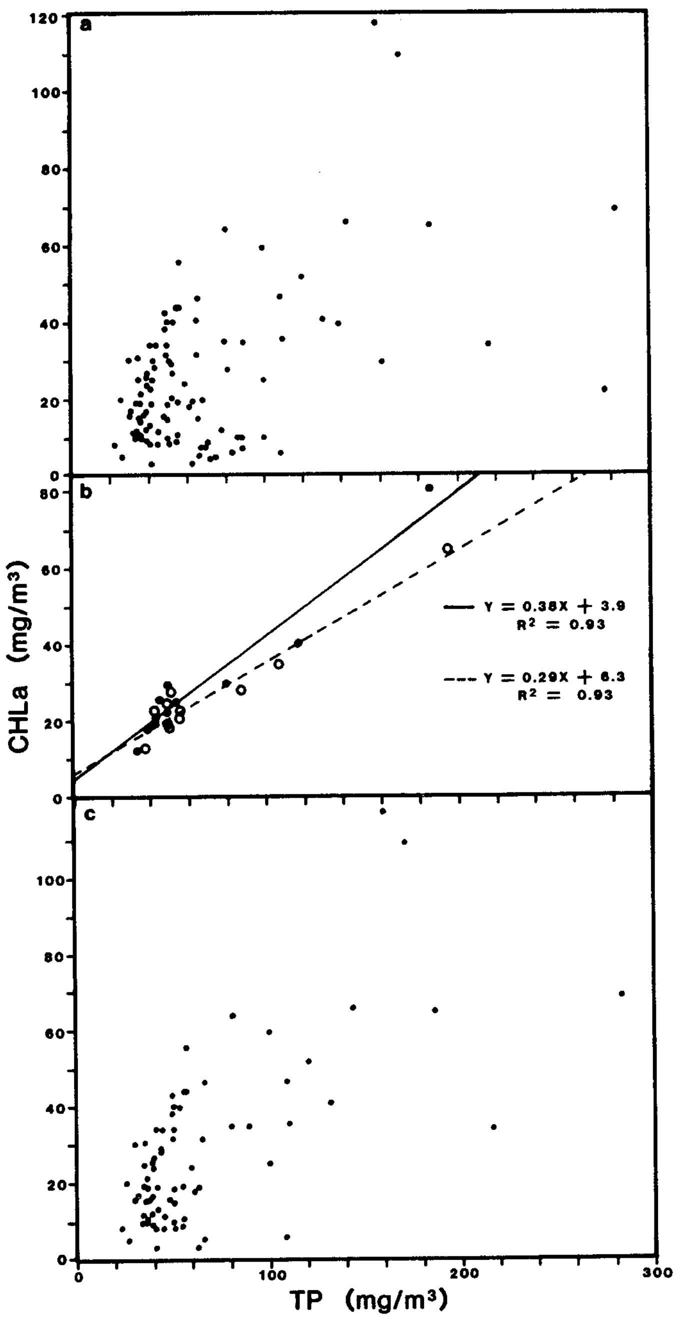

Figure 6. TP and CHLa from 11 Locations on Lake Minnetonka, Minnesota, 1968-1969 (Megard, 1972). A) – Unaveraged TP and CHLa measurements from spring, summer, and fall. B) – Summer (May-September) averages (solid circles) and regression model (solid line); and combined spring, summer, fall averages (open circles) and fitted regression model (dashed line). C) – Unaveraged TP and CHLa measurement from summer (May-September).

Assessment of Variation in TP and CHLa

From these analyses we find that the average precision of summer mean and lake mean values of TP and CHLa obtained in multi-lake surveys is relatively low. The imprecision of these data is a result of the small numbers of measurements (i.e., numbers of samples per summer and numbers of years sampled) typical of lake surveys and the relatively high variability of TP and CHLa over time within individual lakes. For example, although the low average frequency of sampling represented in the present data sets is unlikely to detect the highest or lowest values of TP and CHLa in lakes during a given summer, the average observed ranges of TP and CHLa in these data sets were over 1.8 fold and 2.6 fold, respectively, and fluctuations of this magnitude often occurred in within 1-2 weeks. This short-term, within-lake variation may obscure real differences among water bodies and changes in the long-term condition of individual lakes. The precision of summer means of CHLa based on four samples per summer, for example, will usually be too low to detect (as statistically significant) a real change of two-fold in mean summer CHLa between two years (Figure 5).

While the average precision of lake survey data is low by some standards, there are no widely recognized quantitative criteria for assessing the adequacy of a given level of statistical precision in such data. In addressing this problem we recognize the need for precision varies with the specific uses for the data. In general assessments, such as examining regional patterns of limnology, ranking individual lakes in broad trophic categories, or detecting large changes in the average condition of lakes over time, the precision of data like those described here may be adequate. For example, Vollenweider and Kerekes (1981) have suggested that the general categories of oligotrophy, mesotrophy, and eutrophy are distinguished by approximately three-fold differences in average TP and CHLa. Based on the present analyses, a sampling frequency of four samples per summer would provide about a 99.8 percent probability of detecting a three-fold difference between two TP summer means. The same data, however, would provide only a 50 percent probability of detecting a 1.5-fold difference.

Other potential uses of lake survey data may require more precision. For example, fishery yields of lakes may be linearly related to mean TP and CHLa (Hansen and Leggitt, 1982; Jones and Hoyer, 1982). Differences in TP and CHLa means too small to be distinguished with lake survey data could be of substantial importance in assessing the productivity of lake fisheries. Likewise, studies concerned with measuring small differences in nutrient concentration or algal biomass between lakes or within lakes over time may require greater precision in these measurements that would generally be provided by the lake survey sampling design represented here.

Our present analysis indicates that the precision of lake survey data can be substantially improved by increasing sampling frequency in a particular summer and by increasing the number of years in which data are collected. In studies principally concerned with obtaining information from a large group of lakes, however, such adjustments of effort may not be feasible. Also, it is possible that modifications of sampling protocols other than increasing the frequency of sampling could be more efficient. For example, it is known that large changes in surface water CHLa can result from redistribution of existing phytoplankton as well as from actual changes in the standing crop over time (Fee, 1979; Stauffer, 1981). In

TABLE 6. Regression Models of TP-CHLa Relations Compared Between Data Sets with Observed Temporal Variance (S_d^2) Above and Below the Median Values for All Data Sets.

	Model	R^2	F-Test*
All Data	$\log_{10}CHLa = 1.062 \log_{10}TP - 0.509$	0.63	
Data Divided by TP Variance			
$S_{d-tp}^2 \geqslant 0.0134$**	$\log_{10}CHLa = 0.941 \log_{10}TP - 0.328$	0.55	
			3.86***
$S_{d-tp}^2 < 0.0134$	$\log_{10}CHLa = 1.196 \log_{10}TP - 0.706$	0.72	
Data Divided by CHLa Variance			
$S_{d-chla}^2 \geqslant 0.0357$**	$\log_{10}CHLa = 0.948 \log_{10}TP - 0.343$	0.55	
			5.60***
$S_{d-chla}^2 < 0.0357$	$\log_{10}CHLa = 1.226 \log_{10}TP - 0.740$	0.73	

*General Linear Models Test (Neter and Wasserman, 1974: p. 160).
**Medians for data.
***Significant at $\alpha < 0.05$.

such situations, the precision of estimated summer means might benefit more from the collection of integrated water column samples and from sampling at several widely spaced locations than from increasing the frequency of collecting surface water samples at a single location. Further work in this area is needed to determine the cost-effectiveness of different sampling regimes in increasing the precision of data from large-scale lake surveys.

Lake surveys are part of a class of large-scale limnological studies which have the general goal of identifying and quantifying the major components of variation of limnological conditions within and among lakes (Collins and Sprules, 1983). The major thrust of such research has been to quantify patterns of variation among lakes, by using averages of parameters such as TP and CHLa from water bodies over a broad range of conditions. This approach is taken both from inherent interest in such patterns and because such quantitative relations are used as lake management tools in predicting the effects of large-scale changes within individual lakes. A problem in this approach is that limnological conditions within lakes routinely exhibit a relatively large degree of variation as a result of random or cyclic processes which do not result in any long-term alteration of the lake ecosystem. Our analyses have shown that the magnitude of this short-term, within-lake variability can be large relative to the differences, among or within lakes, which these studies seek to measure. This variation results in uncertainty in measuring the average condition of lakes over time, and reduces our ability to detect differences between lakes and the effects of long-term changes within lakes. In a collection of data from many lakes the range of within-lake variation is likely to be large and may have a considerable influence on the use of such data in quantifying patterns of variation among lakes. The effect of differences in observed temporal variation on the estimation of TP-CHLa regression parameters demonstrated here suggests that further work is needed to clarify the effects of within-lake variation on the use of mean values in predictive models. Within-lake variability is a fundamental quantitative characteristic of many important limnological parameters. The magnitude and effects of this variation needs to be considered more thoroughly in the interpretation and quantitative treatment limnological data.

ACKNOWLEDGMENTS

This research was supported by Missouri Agricultural Experiment Station (Project 176), the Paul K. Wehmiller Fellowship, and the Missouri Cooperative Fisheries Research Unit. We thank James Burroughs, Terry Noonan, Daniel E. Canfield, Jr., Richard Osgood, David Cox, and Mark Kaiser for help in assembling and analyzing the data and for comments on the manuscript.

LITERATURE CITED

Canfield, D. E. and R. W. Bachmann, 1981. Prediction of Total Phosphorus Concentrations, Chlorophyll a, and Secchi Depths in Natural and Artificial Lakes. Can. J. Fish. Aquat. Sci. 38:414-423.

Carlson, R. E., 1977. A Trophic State Index for Lakes. Limnol. Oceanogr. 22:361-369.

Collins, N. C. and W. G. Sprules, 1983. Introduction to Large-Scale Comparative Studies of Lakes. Can. J. Fish. Aquat. Sci. 40:1750-1751.

Edmondson, W. T. and J. T. Lehman, 1981. The Effect of Changes in the Nutrient Income on the Condition of Lake Washington. Limnol. Oceanogr. 26:1-29.

Fee, E. J., 1976. The Vertical and Seasonal Distribution of Chlorophyll in Lakes of the Experimental Lakes Area, Northwestern Ontario: Implications for Primary Production Estimates. Limnol. Oceanogr. 21:767-783.

Hanson, J. M. and W. C. Leggett, 1982. Empirical Prediction of Fish Biomass and Yield. Can. J. Fish. Aquat. Sci. 39:257-263.

Hoyer, M. V. and J. R. Jones, 1983. Factors Affecting the Relation Between Phosphorus and Chlorophyll a in Midwestern Reservoirs. Can. J. Fish. Aquat. Sci. 40:192-199.

Jones, J. R. and R. W. Bachmann, 1976. Prediction of Phosphorus and Chlorophyll Levels in Lakes. J. Water Pollut. Control Fed. 48:2176-2182.

Jones, J. R. and R. W. Bachmann, 1978a. Trophic Status of Iowa Lakes in Relation to Origin and Glacial Geology. Hydrobiologia 57:267-273.

Jones, J. R. and R. W. Bachmann, 1978b. Phosphorus Removal by Sedimentation in Some Iowa Reservoirs. Verh. Internat. Verin. Limnol. 20:1576-1580.

Jones, J. R. and M. V. Hoyer, 1982. Sportfish Harvest Predicted by Summer Chlorophyll-a Concentration in Midwestern Lakes and Reservoirs. Trans. Am. Fish. Soc. 111:176-179.

Jones, J. R. and J. T. Novak, 1981. Limnological Characteristics of Lake of the Ozarks, Missouri. Verh. Internat. Verin. Limnol. 21: 919-925.

Jones, R. A., W. Rast, and G. F. Lee, 1979. Relationship Between Summer Mean and Maximum Chlorophyll a Concentrations in Lakes. Environm. Sci. Tech. 13:869-870.

Kalff, J. and R. Knoechel, 1978. Phytoplankton and Their Dynamics in Oligotrophic and Eutrophic Lakes. Ann. Rev. Ecol. Sys. 9:475-495.

Likens, G. E., 1975. Primary Productivity of Inland Aquatic Ecosystems. *In:* Primary Productivity of the Biosphere, H. Lieth and E. H. Whittaker (Editors). Springer-Verlag, New York, New York, pp. 185-202.

Lillie, R. A. and J. W. Mason, 1983. Limnological Characteristics of Wisconsin Lakes. Technical Bulletin No. 138, Wisconsin Dept. of Natural Resources, Madison, Wisconsin.

Megard, R. O., 1972. Phytoplankton, Photosynthesis, and Phosphorus in Lake Minnetonka, Minnesota. Limnol. Oceanogr. 17:68-87.

Moyle, J. B., 1956. Relationships Between the Chemistry of Minnesota Surface Waters and Wildlife Management. J. Wildl. Mgmt. 20: 303-320.

Neter, J. and W. Wasserman, 1974. Applied Linear Statistical Models. Richard D. Irwin, Inc., Homewood, Illinois, 842 pp.

Noonan, T. A., 1979. Crustacean Zooplankton and Chlorophyll a Relationships in Some Iowa Lake and Reservoirs. Master's thesis, Library Iowa State University, Ames, Iowa.

Oglesby, R. T., 1977. Relationships of Fish Yield to Lake Phytoplankton Standing Crop, Production, and Morphoedaphic Factors. J. Fish. Res. Board Can. 34:2271-2279.

Osgood, R. A., 1981. A Study of the Water Quality of 60 Lakes in the Seven County Metropolitan Area. Metropolitan Council Publ. No. 01-81-047, St. Paul, Minnesota.

Osgood, R. A., 1982. A 1981 Study of the Water Quality of 30 Lakes in the Seven-County Metropolitan Area. Metropolitan Council Publication No. 10-82-005, St. Paul, Minnesota.

Prepas, E. E. and F. H. Rigler, 1982. Improvements in Quantifying the Phosphorus Concentration in Lake Water. Can. J. Fish. Aquat. Sci. 39:822-829.

Reckhow, K. H., 1979. Quantitative Techniques for the Assessment of Lake Quality. U.S. Environmental Protection Agency Report No. EPA-440/5-79-015.

Reckhow, K. H. and S. C. Chapra, 1983. Confirmation of Water Quality Models. Ecol. Modelling 20:113-133.

Sakamoto, M., 1966. Primary Production of Phytoplankton Community in Some Japanese Lakes and Its Dependence on Lake Depth. Arch. Hydrobiol. 62:1-28.

Shapiro, J. and H. Pfannkuch, 1974. The Minnesota Chain of Lakes. A Study of Urban Drainage and Its Effects. Interim Report No. 9, Limnological Research Center, University of Minnesota, Minneapolis, Minnesota.

Smith, V. H., 1979. Nutrient Dependence of Primary Productivity in Lakes. Limnol. Oceanogr. 24:1051-1064.

Smith, V. H. and J. Shapiro, 1981. Chlorophyll-Phosphorus Relations in Individual Lakes. Their Importance to Lake Restoration Strategies. Environm. Sci. Tech. 15:444-451.

Snedecor, G. W. and W. G. Cochran, 1980. Statistical Methods. Seventh Edition. The Iowa State University Press, Ames, Iowa, 507 pp.

Statistical Analysis System, 1982. SAS Users Guide. SAS Institute, Inc., Raleigh, North Carolina.

Stauffer, R. E., 1981. Sampling Strategies for Estimating Chlorophyll Standing Crops in Stratified Lakes. *In:* Restoration of Lakes and Inland Waters. U.S. Environmental Protection Agency, E.P.A. 400/5-81-010, pp. 203-209.

Stauffer, R. E., G. F. Lee, and D. E. Armstrong, 1979. Estimating Chlorophyll Extraction Biases. J. Fish. Res. Board Can. 36:152-157.

Thornton, K. W., R. H. Kennedy, A. D. Magoun, and G. E. Saul, 1982. Reservoir Water Quality Sampling Design. Water Resources Bulletin 18(3):471-480.

Vollenweider, R. A., 1975. Input-Output Models with Special Reference to the Phosphorus Loading Concept in Limnology. Schweiz. Zeit. Hydrol. 37:53-84.

Vollenweider, R. A. and J. J. Kerekes, 1981. Background and Summary Results of the OEDC Cooperative Program on Eutrophication. *In:* Restoration of Lakes and Inland Waters. U.S. Environmental Protection Agency, E.P.A. 400/5-81-010, pp. 25-36.

Walker, W. W., Jr., 1979. Use of Hypolimnetic Oxygen Depletion Rate as a Trophic State Index for Lakes. Water Resources Res. 15:1463-1469.

Walmsley, R. D., 1984. A Chlorophyll a Trophic Status Classification System for South African Impoundments. J. Environ. Qual. 13:97-104.