United States
Environmental Protection
Agency

Office of Water
Regulations and Standards
Washington, D.C. 20460

EPA 440/5-81-010

Water

EPA

RESTORATION OF LAKES AND INLAND WATERS

VARIABILITY OF TROPHIC STATE INDICATORS IN RESERVOIRS

WILLIAM W. WALKER, JR.
Environmental Engineer
Concord, Massachusetts

ABSTRACT

As part of the Environmental Water Quality Operation Studies being conducted by the Army Corps of Engineers, a data base has been compiled that describes the morphometry, hydrology, and water quality of over 300 reservoirs throughout the United States. The data base will be used to test and evaluate existing empirical models for assessing eutrophication problems and to develop new methods, where appropriate. This work has been motivated by concerns over the application of existing models to reservoirs, despite the fact that most have been developed using data bases consisting entirely of northern, natural lakes. Existing methods may not be adequate for reservoirs because of differences in morphometry, hydraulics, sedimentation, and region, that may influence responses to nutrient loading. To provide preliminary insights into the effects of using different data-reduction procedures and into the adequacy of the data for model testing purposes, EPA National Eutrophication Survey data from 76 phosphorus-limited Corps impoundments are analyzed and used in testing Carlson's (1977) Trophic State Indices. Seasonal effects and variance/covariance components are identified at different averaging levels. Results indicate that chlorophyll *a* levels in Corps reservoirs are generally less sensitive to phosphorus or transparency than in the natural lakes used by Carlson in developing the index system. The use of error analysis for assessing the adequacy of the data set for model testing purposes is demonstrated.

INTRODUCTION

The development of phosphorus loading/trophic state response models over the past decade has greatly increased the feasibility of lake water quality planning. Most of these models have been based upon empirical studies of data from natural lakes in glaciated regions. Their applicability to manmade impoundments is in question because of lake/reservoir differences in age, morphometry, hydrodynamics, sedimentation, and region (Thornton, et al. 1980). To provide a basis for testing available models, data describing the morphometry, hydrology, water quality, and sedimentation rates of over 300 active U.S. Army Engineer reservoirs have been compiled (Walker, 1980a). During the next year, this data base will be used in a systematic assessment of phosphorus loading models and relationships among trophic state indicators in reservoirs.

The data base currently contains over two million water quality observations. Testing empirical eutrophication models in reservoirs requires averaging water quality measurements over spatial and temporal scales. If within-pool water quality variations are not random with respect to date, station, or depth, then summary statistics for a given reservoir will depend to some extent upon the particular data reduction method employed. The choice of reduction method may, in turn, influence conclusions regarding the adequacy of existing models as well as the parameter estimates of any new models which may be developed.

There is no standard data reduction procedure which can be used prior to model development, testing, or application. Methods have included, for example, (1) taking the median or mean of all within-pool observations (U.S. EPA, 1975); (2) sequential averagi over depths, stations, and dates (Lambou, et al. 197 (3) sequential averaging within specific depth rang (Carlson, 1977); and (4) various weighted averagi schemes which reflect morphometric characteristi (Boyce, 1973). As compared with natural lakes, ma reservoirs pose special data reduction proble because of extreme spatial and/or temporal variatio in conditions.

This paper describes investigations of the variabil of trophic state indicators among and within a group Corps reservoirs. The analysis covers seasonal r lationships, variance/covariance components, regre sion analyses, and error analyses. This work has be undertaken to assess the implications of using differe data reduction procedures and to assess the adequa of the data for model testing purposes.

DATA BASE

National Eutrophication Survey (U.S. EPA, 197 data have been used as a basis for this analysis. Th relatively uniform sampling program designs used b the survey provide data that are suitable for statistic treatment. One drawback, however, is that under th program reservoirs were typically sampled only thre times during one growing season. In future work, w plan to examine data from other agencies, which, many cases, are more intensive and/or cover long periods. The Survey data have been screened t eliminate data from 19 reservoirs which we predominately nitrogen-limited (based upon bioassay and to eliminate all stations with fewer than thre sampling dates for total phosphorus, chlorophyll *a*, an transparency. The resulting file contains 963 observations from 306 stations in 76 reservoirs.

Surface total phosphorus, Secchi depth, and chlorophyll *a* values have been expressed in terms of Carlson's Trophic State Indices (Carlson, 1977):

$$I_P = 4.2 + 33.2 \log_{10} P \quad \text{eq. 1}$$
$$I_T = 60 - 33.2 \log_{10} Z_s \quad \text{eq. 2}$$
$$I_B = 30.6 + 22.6 \log_{10} B \quad \text{eq. 3}$$

where,

P = total phosphorus concentration (mg/m^3)
Z_s = Secchi depth (m)
B = chlorophyll *a* concentration (mg/m^3)
T = transparency

The indices are calibrated so that the three versions are equivalent, on the average, when applied to midsummer, epilimnetic data from northern, natural lakes. Expression of measurements on these scales tends to reduce the skewness in the distributions of the variables and provides benchmarks for assessing reservoir trophic state relationships in comparison to those typical of natural lakes.

The latitudes of 309 natural lakes sampled by the Survey are compared with the latitudes of 106 Corps reservoirs sampled by the Survey in Figure 1. The distribution of natural lakes is bimodal, with a northern peak (glacial lakes) and a southern peak (subtropical lakes in Florida). Most of the Corps reservoirs may be influenced by regional factors as well as the effects of impoundment type.

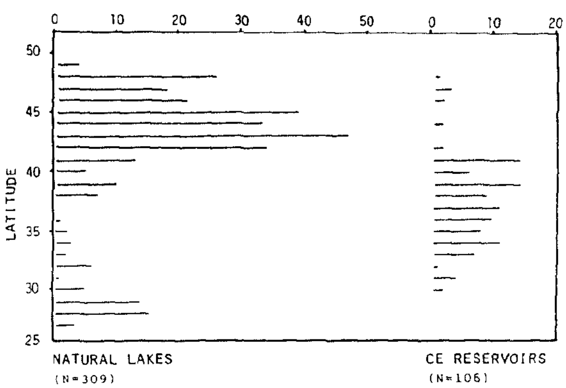

Figure 1. — Latitudes of natural lakes and Corps reservoirs sampled by the EPA National Eutrophication Survey.

SEASONAL RELATIONSHIPS

Average seasonal variations in the index components are depicted in Figure 2. Station means have been computed and their effects removed from the data prior to calculating the mean and standard error for each month (March to November) and index component. Analyses of variance indicate that fixed monthly effects are significant ($p < .0001$) but explain only 11 percent of the total within-station variance of each index. The seasonal variations depicted in Figure 2 are characteristic of this collection of reservoirs but not necessarily of each individual reservoir.

Average seasonal effects on phosphorus and transparency are similar: Both tend to be lowest during March and midsummer and highest during April and November, possibly reflecting seasonal flow and turbidity variations and the influences of turnover periods. Monthly effects on chlorophyll *a* suggest a spring maximum (April-May), followed by a June depression, a midsummer maximum, and lower values in November. Temperature and light effects may be responsible for the relatively low chlorophyll *a* levels during March and November. The June depression may be caused by seasonal succession of algal species. A more detailed examination of the data indicates that lower June chlorophyll *a* levels are characteristic of about half of the stations sampled in June, while the rest have June levels more typical of May or July values. In testing seasonal aspects of TSI behavior, Carlson (1977) also noted a June depression in chlorophyll *a* index relative to the phosphorus index in three natural lakes.

Differences among various versions of the index provide a measure of "lake-like" behavior, since the index system is calibrated so that I_P, I_T, and I_B, values are equivalent, on the average, when applied to midsummer epilimnetic data from northern, natural lakes. Figure 2 indicates that the range of index means is generally lowest during midsummer and highest during March, June, and November (approaching 15). Minor recalibration of the phosphorus and/or transparency index would bring I and I into agreement for all seasons, since the monthly effect curves in Figure 2 are roughly parallel. Since seasonal chlorophyll *a* behavior is fundamentally different, however, recalibration alone would not eliminate biases (i.e., significant differences between I_B and I_P or I_T) for all seasons.

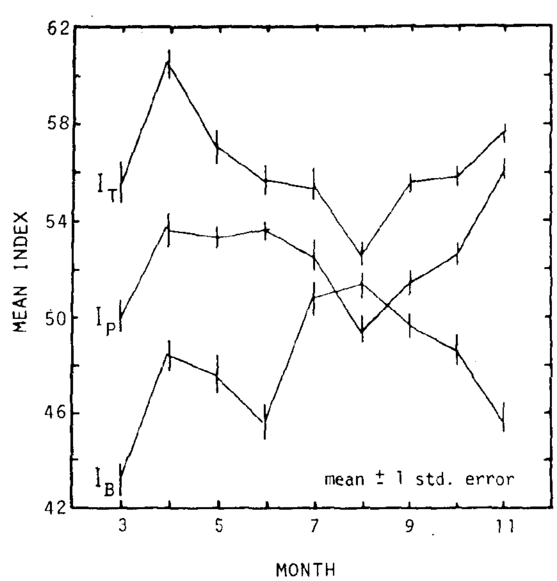

Figure 2. — Monthly variations in trophic state indices.

VARIANCE COMPONENTS

Trophic index observations can be classified in a hierarchy defined by region, reservoir, station, and sampling date. Variations at each level could account for some portion of the total variance of each index. A nested analysis of variance procedure (Statist. Anal. Inst. 1979) has been applied to derive pooled estimates of variance and covariance components according to the following model:

$$\mathrm{Var}(I) = \sigma_d^2 + \sigma_{r(d)}^2 + \sigma_{s(d,r)}^2 + \sigma_e^2 \qquad \text{eq. 4}$$

where,

σ_d^2 = variance among regions, defined by Corps districts

$\sigma_{r(d)}^2$ = variance among reservoirs, within districts

$\sigma_{s(d,r)}^2$ = variance among stations, within reservoirs and districts

σ_e^2 = variance within stations

This model has been used to describe variations in the data. It is of limited use for significance testing, which would require randomness and serial independence in the within-station variations, that can be attributed to variations in time, sampling error, and measurement error. As demonstrated in the previous section, some of the within-station variations can be attributed to seasonal factors and are therefore nonrandom. Given three observations per station spaced at roughly bimonthly intervals, serial dependence in the observations is not likely to be strong, since conditions are known to vary in many reservoirs at a much higher frequency, as influenced, for example, by storm events and algal bloom occurrences. Among-station, within-reservoir variations also show some serial dependence, since spatial trends in the indices are often apparent when station means are displayed in a downstream order (Walker, 1980b).

The relative magnitude of the last term is of special significance to modelling efforts. With relatively large within-station variance, it would be difficult to obtain much accuracy in station summary statistics (e.g., station mean) with limited data. This would reduce the explainable variance of any model or index system calibrated to the reduced data set, make it more difficult to distinguish among alternative model formulations, and increase the error associated with model parameter estimates.

Variance components estimated for each index are displayed on the left side of Figure 3. Variations in the phosphorus and transparency indices are similar at all levels. Variance components of the chlorophyll *a* index at the district, reservoir, and station levels are considerably lower than would be predicted based upon corresponding phosphorus and transparency variance components. The within-station components account for a major portion (~60 percent) of the total chlorophyll *a* variability. Thus, on the average for this data set, temporal variations in the chlorophyll *a* index at a given station appear to be stronger than variations among stations, reservoirs, and/or districts. The within-station variance components correspond to standard deviations of 6.5, 6.5 and 7.9 for I_B, I_T, and I_P respectively.

The covariance components on the right side c Figure 3 provide insights into relationships among th indices at different averaging levels. Spatial cc variances are positive in all cases. Thus, the variou versions of the index correlate positively amon districts, reservoirs within districts, and stations withi reservoirs. Appreciable temporal covariance is ob served only for phosphorus and transparency. Thi covariance might be attributed, for example, to turbidit variations following seasonal or short-term (storr event) flow variations. Despite its positive covarianc spatially, the chlorophyll *a* index does covary ter porally with the other indices.

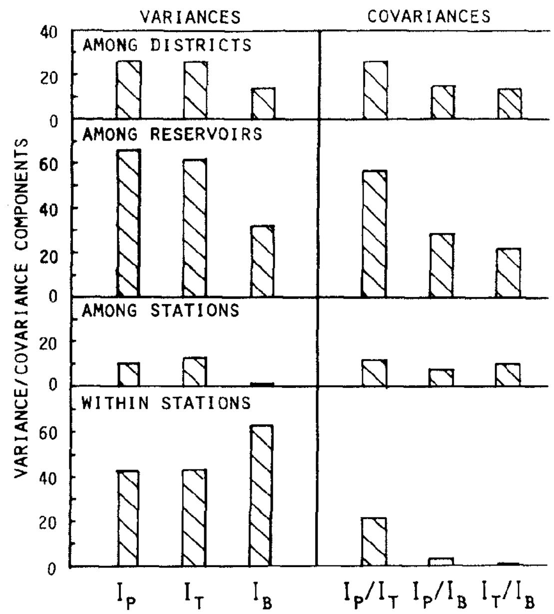

Figure 3. — Variance/covariance components of trophic st indices.

REGRESSION ANALYSES

The covariance components indicate that the indic can be averaged by station with some loss information about the phosphorus/transparency r lationship, but without losing chlorophyll *a* predic ability. Figure 4 depicts relationships among statio mean values of the indices. Results of standard a geometric-mean regression analyses relating t indices are given in Table 1. Geometric me regressions summarize functional relationships a are appropriate to use when both the independent a dependent variables are subject to natural variabil and measurement error (Ricker, 1973). Standa regressions are appropriate for predictive purposes

The phosphorus and transparency indices explain and 29 percent of the variance in the chlorophyll index, respectively. Recalibration of the index system these reservoirs requires significant reductions in t corresponding slopes. In contrast, the phosphor index explains 78 percent of the transparency ind

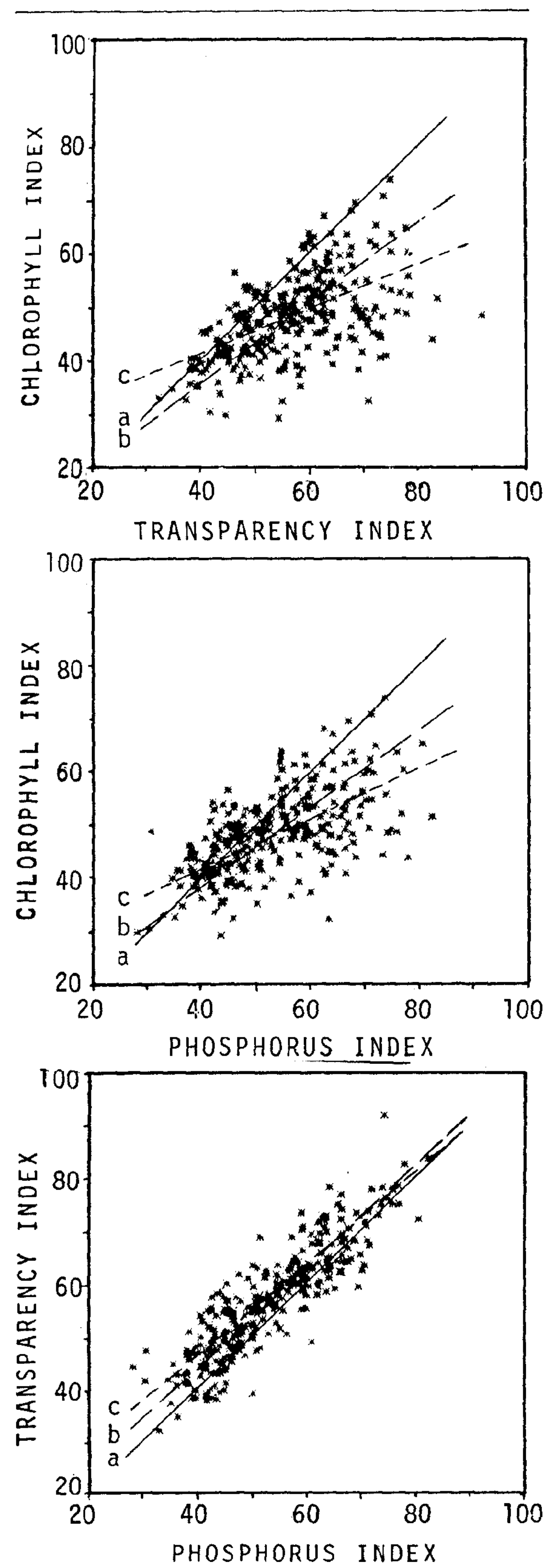

Figure 4. — Relationships among station-mean index values (a = line of equality, b = geometric-mean regression, c = standard regression).

variance and requires an adjustment in the intercept only. Thus, compared with chlorophyll *a*, the phosphorus/transparency relationship appears to be more typical of the natural lakes used by Carlson in deriving his index system.

The effects of using alternative data reduction procedures on the regression analyses have been also studied. Using only summer mean values reduces the regression slopes and R^2 values and increases mean squared residual errors by 58, 46, and 94 percent for the I_B/I_P, I_B/I_T, and I_T/I_P regressions, respectively. These increases in error result partially from loss of within-station replication when spring and fall values are eliminated. In future work, data from other monitoring programs with more intensive summer sampling will be investigated. Use of reservoir means has little influence on the results, but increases the standard errors of parameter estimates.

ERROR ANALYSES

Residual errors from the regressions can be attributed to three types of error: parameter, data, and model. The first reflects uncertainty in the model coefficients; the second, errors in the predicted and/or predictor variables; and the third, influences of factors which are not considered in the model structure. The results of the variance component and regression analyses can be used to derive approximate estimates of the data errors according to the following equation:

$$Var(R)_D \quad \frac{Var(I_Y)_E + b^2 Var(I_X)_E - 2b\, Cov(I_Y, I_X)_E}{N} \qquad \text{eq. 5}$$

where,

$Var(R)_D$ = data-error component of mean-squared residual
$Var(I_Y)_E$ = within-station variance of predicted index
$Var(I_X)_E$ = within-station variance of predictor index
b = slope of regression equation
$Cov(I_Y, I_X)_E$ = within-station covariance of predicted and predictor indices
N = number of observations per station (averaging 3.1)

This formula is approximate because it assumes serial independence in the within-station variations, which would tend to be more important at sampling intervals less than the 2-month intervals characteristic of this data set.

The results of applying this equation to the regression models in Table 1 are given in Table 2. They indicate that roughly half (50 to 59 percent) of the residual errors from the regressions can be attributed to data errors. These components could be reduced with a more intensive sampling program (i.e, more replications per station). The influences of parameter uncertainty on the total residual error are expected to be relatively insignificant, since the parameter error component is inversely proportional to the number of stations used in the regression analyses and the parameters are relatively well-determined (Walker, 1977). Thus, most of the remaining error can be attributed to the effects of factors which are not considered in the index system.

Since data errors do not explain all of the residual variance, it may be possible to improve the index system by modifying it to take other important factors into account. One modification is suggested by these results and by the turbid nature of many reservoirs. Chlorophyll *a*/phosphorus and chlorophyll *a*/transparency relationships may not be constant across reservoirs because of variations in non-algal particulate materials (turbidity), which would influence measurements of total phosphorus and transparency but not of chlorophyll *a*. The relative stability of the phosphorus/tranparency relationship across lakes and reservoirs may be attributed to the fact that both types of measurements are sensitive to algal and non-algal particulate materials. Other factors which might contribute to model error include kinetic effects in reservoirs with short hydraulic residence times. It might also be possible to modify the system to account for nitrogen limitation, by including N-limited as well as P-limited reservoirs in the data set. Expansion of the index system to include hypolimnetic oxygen deficits is another possibility (Walker, 1979). These approaches will be investigated in future studies of Carlson's index system and other schemes using more extensive and intensive data sets derived from the Corps reservoir data base.

CONCLUSIONS

This paper has demonstrated an analytical approach which provides insights into the adequacy of data for modeling purposes. Potential applications of the approach to monitoring program design are discussed elsewhere (Walker, 1980b). Results suggest that chlorophyll *a* is considerably less sensitive to phosphorus or transparency in these reservoirs, compared with the natural lakes used by Carlson in developing the index system. The phosphorus and transparency indices are not relative indicators of biomass in these reservoirs, possibly because they are influenced b non-algal materials. In future work the approache demonstrated here will be applied in evaluatin alternative schemes for summarizing relationshi among measures of trophic state in reservoirs, usir an expanded data set.

Table 1. — Results of regression analyses relating station-mean index values.

Equation	Slope standard error	R^2	Mean squared error
	standard regressions		
$I_B = 24.7 + .443\, I_P$	.033	.374	37.2
$I_B = 25.2 + .403\, I_T$	.036	.291	42.1
$I_T = 11.6 + .854\, I_P$	.026	.774	24.1
	geometric mean regressions		
$I_B = 9.9 + .724\, I_P$	.033	.224	46.1
$I_B = 5.7 + .747\, I_T$	.036	.079	54.7
$I_T = 5.4 + .971\, I_P$	.026	.760	25.6

Table 2. — Results of error analyses.

Relationship*	Mean squared error	Data error	Percent data error
I_B/I_P	37.2	21.8	59%
I_B/I_T	42.1	22.0	52%
I_T/I_P	24.1	12.1	50%

* Standard regressions in Table 1.

REFERENCES

Boyce, F. M. 1973. A computer routine for calculating to lake volume contents of a dissolved substance from arbitrary distribution of concentration profiles. Tech. Bı 83, Inland Waters Directorate, Canada Centre for Inla Waters, Burlington, Ontario.

Carlson, R. E. 1977. A trophic state index for lakes. Limn Oceanogr. 22:361.

Lambou, V. W., et al. 1976. Prediction of phytoplank productivity in lakes. *In* W. R. Ott, ed. Environmer modeling and simulation. Off. Res. Dev. Off. Plann. Mana U.S. Environ. Prot. Agency.

Ricker, W. E. 1973. Linear regressions in fishery resea Jour. Fish. Res. Board Can. 30:409.

Statistical Analysis Institute. 1979. SAS User's Guide.

Thornton, K. W., et al. 1980. Reservoir sedimentation water quality — a heuristic model. Surface w impoundments. Am. Soc. Civil. Eng. and University Minnesota, Minneapolis.

U.S. Environmental Protection Agency. 1973-1976. Nati Eutrophication Survey Working Pap. Ser., Pacific N Environ. Res. Lab., Corvallis, Ore.

________. 1975. National Eutrophication Survey Meth 1973-76. Working Pap. No. 175. Pacific N.W. Environ. Lab., Corvallis, Ore.

Walker, W. W. 1977. Some analytical methods applied to water quality problems. Ph.D. thesis. Harvard Univer

________. 1979. Use of hypolimnetic oxygen depletion as a trophic state index for lakes. Water Resour. 15:1463.

________. 1980a. Empirical methods for predicting trophication problems in impoundments — phase I base development. U.S. Army Corps Eng. Waterways Sta. Vicksburg, Miss. Draft final rep.

________. 1980b. Analysis of water quality variatior reservoirs: Implications for monitoring and mode efforts. Symp. Surface Water Impoundments. Am. Soc. Eng. and University of Minnestoa, Minneapolis.