W. W. WALKER
1127 Lowell Road
Concord, MA. 01742

Lake and Reservoir Management- *Practical Applications*

Water Quality Criteria and Standards

STATISTICAL BASES FOR MEAN CHLOROPHYLL *A* CRITERIA

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

ABSTRACT

Mean chlorophyll *a* concentration has been frequently used in trophic state classification. Recently, increased attention has been given to extreme conditions (such as maximum chlorophyll *a* or nuisance bloom frequency) as important descriptors because they have the greatest potential for directly affecting water uses. Frequency distribution models for representing chlorophyll *a* temporal variability are calibrated to three independent data sets and used to predict relationships between mean chlorophyll *a* and percent of time various extreme values (such as 20, 30, 40 ppb) are exceeded. These extreme levels have been associated with bloom or nuisance-level conditions. The responses are nonlinear and suggest a threshold arithmetic-mean chlorophyll *a* of approximately 10 ppb, below which expected bloom frequencies are minimal. Bloom frequencies rise sharply as mean chlorophyll *a* increases from 10 ppb. Results are largely independent of the particular variance model employed. The statistically-defined threshold value of mean chlorophyll *a* corresponding to the onset of detectable bloom frequencies agrees with subjective definitions of the mesoeutrophic boundary. The methodology established here has potential applications in formulating lake water quality objectives.

INTRODUCTION

Chlorophyll *a* is the most direct and practical measurement of algal productivity and eutrophication response in impoundments. Several studies have related chlorophyll *a* or other measures of algal standing crop to water quality aspects which directly impact water uses, including transparency, hypolimnetic oxygen depletion, fish production, taste-and-odor episodes, blue-green toxicity, and trihalomethane precursors.

This paper demonstrates a conceptual framework for relating mean chlorophyll *a* values to indices of use impairment. By employing statistical frequency distributions, the effects of temporal variability in chlorophyll on use impairment are incorporated. The framework could be used in developing regional or site-specific criteria or standards for protecting lake and reservoir uses from eutrophication-related impacts (Walker, 1984).

Table 1.—Definitions of the mesotrophic/eutrophic boundary.

Mean Chl. *a* (ppb)	Source
10	National Academy of Science (1972)
15	Sakamoto (1966)
8.8	Dobson (1974)
12	USEPA National Eutrophication Survey (1974)

Several definitions of the mesotrophic/eutrophic boundary based upon mean chlorophyll *a* values are presented in Table 1. These values were subjectively derived by various investigators. Definitions of *eutrophic* are of mixed origin and do not necessarily correspond with levels of use impairment. Use of mean chlorophyll *a* values as relative measures of lake condition has certain advantages, including:

Water Quality Criteria and Standards

STATISTICAL BASES FOR MEAN CHLOROPHYLL *A* CRITERIA

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

ABSTRACT

Mean chlorophyll *a* concentration has been frequently used in trophic state classification. Recently, increased attention has been given to extreme conditions (such as maximum chlorophyll *a* or nuisance bloom frequency) as important descriptors because they have the greatest potential for directly affecting water uses. Frequency distribution models for representing chlorophyll *a* temporal variability are calibrated to three independent data sets and used to predict relationships between mean chlorophyll *a* and percent of time various extreme values (such as 20, 30, 40 ppb) are exceeded. These extreme levels have been associated with bloom or nuisance-level conditions. The responses are nonlinear and suggest a threshold arithmetic-mean chlorophyll *a* of approximately 10 ppb, below which expected bloom frequencies are minimal. Bloom frequencies rise sharply as mean chlorophyll *a* increases from 10 ppb. Results are largely independent of the particular variance model employed. The statistically-defined threshold value of mean chlorophyll *a* corresponding to the onset of detectable bloom frequencies agrees with subjective definitions of the mesoeutrophic boundary. The methodology established here has potential applications in formulating lake water quality objectives.

INTRODUCTION

Chlorophyll *a* is the most direct and practical measurement of algal productivity and eutrophication response in impoundments. Several studies have related chlorophyll *a* or other measures of algal standing crop to water quality aspects which directly impact water uses, including transparency, hypolimnetic oxygen depletion, fish production, taste-and-odor episodes, blue-green toxicity, and trihalomethane precursors.

This paper demonstrates a conceptual framework for relating mean chlorophyll *a* values to indices of use impairment. By employing statistical frequency distributions, the effects of temporal variability in chlorophyll on use impairment are incorporated. The framework could be used in developing regional or site-specific criteria or standards for protecting lake and reservoir uses from eutrophication-related impacts (Walker, 1984).

Table 1.—Definitions of the mesotrophic/eutrophic boundary.

Mean Chl. *a* (ppb)	Source
10	National Academy of Science (1972)
15	Sakamoto (1966)
8.8	Dobson (1974)
12	USEPA National Eutrophication Survey (1974)

Several definitions of the mesotrophic/eutrophic boundary based upon mean chlorophyll *a* values are presented in Table 1. These values were subjectively derived by various investigators. Definitions of *eutrophic* are of mixed origin and do not necessarily correspond with levels of use impairment. Use of mean chlorophyll *a* values as relative measures of lake condition has certain advantages, including:

1. Estimates of mean (or median) values derived from a given monitoring program would generally have lower variance (that is, be more reliably estimated from limited data) than other summary statistics, such as maximum;
2. Mean chlorophyll *a* can be related to watershed conditions using nutrient budget models; and
3. Mean values have been widely used in lake assessment and classification.

Temporal variations in chlorophyll *a* within a given waterbody and growing season can be substantial and are generally not reflected by a single mean value. The Organization for Economic Cooperative Development (OECD) (1982) eutrophication study found that maximum chlorophyll *a* concentrations averaged 3.14 times the annual mean values. This temporal variability is too easily overlooked in formulating management objectives, particularly when decisionmakers are presented only with mean concentrations or trophic state indices.

Recently, increased attention has been given to extreme conditions (maximum chlorophyll *a* or nuisance-level frequency) as lake condition indices (OECD, 1982; Walmsley, 1984). One problem with using maximum chlorophyll *a* in a classification system is that the maximum value detected in a monitoring program would depend upon the number of samples taken. For example, for a given station and growing season, a weekly monitoring program would, on the average, detect a higher maximum concentration than a monthly monitoring program. Use of the 95th percentile (instead of the absolute maximum) would eliminate the above dependence, but would require intensive sampling schedules for reliable direct measurement in each water body. Perhaps a more practical approach would be to infer extreme values (95th percentiles) by fitting statistical frequency distributions to monitoring data.

The framework developed here is based upon the hypothesis that water use impacts (aesthetics, recreation, and water supply) are more directly related to instantaneous chlorophyll *a* concentrations than to annual or seasonal mean values. A swimmer who encounters a floating algal mat on Saturday is not really consoled by the fact that the water was clear on Tuesday. While mean or median conditions may be acceptable in a water supply, taste-and-odor episodes, treatment-plant upsets, and trihalomethane violations may occur during or following intermittent algal blooms (Walker, 1983a; Bernhardt, 1984).

Statistical frequency distribution models are calibrated for describing chlorophyll *a* variability at a given station and growing season. By mapping models which relate use impairment to instantaneous concentrations onto the chlorophyll *a* frequency distribution, it is possible to develop indices of use impact which are sensitive to temporal variability but which can be predicted from mean chlorophyll *a* concentrations.

USE IMPACT MODELS

Impacts of eutrophication, as measured by chlorophyll *a* concentration, depend upon the types and intensities of water use, regional factors (user adaptation), and dominant algal species. The literature is generally lacking in systematic studies designed specifically for developing quantitative relationships between instantaneous algal concentrations and water uses. Results of a study of 21 South African reservoirs (Walmsley and Butty, 1979; Walmsley, 1984) are used in the following chart in formulating a use impact model. Alternative models could be developed from regional information using this approach.

The South African study involved simultaneous collection of data on water quality (chlorophyll *a*, nutrients, transparency, oxygen profiles, etc.), aesthetics (water appearance, surface scums), and evidence of use impacts (based primarily upon interviews with recreational area managers and water treatment plant operators). Based upon review of their data, Walmsley and Butty ascribed "nuisance values" to certain instantaneous chlorophyll *a* ranges, according to the following scheme:

Chl. *a* (ppb)	Nuisance Value
0–10	*No problems encountered*
10–20	*Algal scums evident*
20–30	*Nuisance conditions encountered*
>30	Severe nuisance conditions encountered

Walmsley (1984) employed these categories in developing a trophic state classification system for South African impoundments. According to the OECD (1982) trophic state classification scheme, an impoundment with a mean chlorophyll *a* concentration exceeding 25 ppb would have greater than a 50 percent probability of being classified hypereutrophic. There appears to be some consensus that instantaneous chlorophyll *a* concentrations exceeding the 20–30 ppb range pose problems for water users.

Four types of functions might be used to relate concentration to use impact (Fig. 1):

(1) Linear–Zero Intercept
(2) Linear–Positive Intercept
(3) Step Function
(4) Multiobjective

The first model assumes that the impact is simply proportional to chlorophyll *a* concentration. This model does not seem appropriate on the basis that algal impacts would tend to be undetectable in the low range of concentration (0–10 ppb according to Walmsley and Butty, 1979). The second model provides an intercept and seems more reasonable on the basis of the South African study. Problems exist, however, in estimating a slope (and possible upper asymptote) for the line, which amount to ascribing specific numeric values to represent the relative impact of nuisance category (for instance, are "severe nuisance" conditions two, four, or ten times as bad as "nuisance" conditions?).

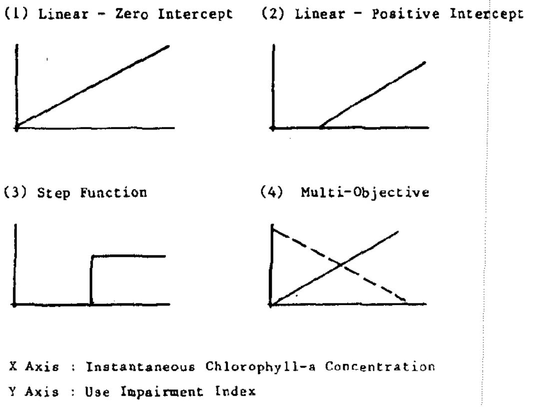

Figure 1.—Hypothetical models for use impairment as a function of instantaneous chlorophyll *a* concentration.

The third model is a step function which assumes that impact is zero below a certain level and one above that level. While this model is highly simplified, it is not unreasonable and has certain advantages. The fourth model represents multiple management objectives. As described by Wagner (1984), higher chlorophyll *a* concentrations may be viewed as beneficial from the standpoint of fishery production under certain conditions (such as, in extremely oligotrophic waters, or also in enriched waters, provided that fishermen are not particular about the species of fish they catch). It seems reasonable that any beneficial effects on fishery production would be more directly related to mean values than to instantaneous values. While this type of model is not being considered in this paper, the methodology (with some increase in complexity) could be applied in situations where multiple objectives and tradeoffs must be considered.

The simplest measure of user impact would be the percent of the time over the algal growth (or recreational) season that nuisance-level conditions are encountered. This measure of impact corresponds to the third model (step function) (Fig. 1). Presumably, this statistic would be related to the frequency of algae-related problems detected by recreational users or water treatment plant operators. An advantage of this expression of impact is that it is more easily grasped by the public and decisionmakers than the expressions of concentration, trophic index, or multivariate index.

In the case of a water treatment plant, a step-function impact model is not entirely unrealistic, since special measures to control algal problems (for example, copper sulfate, powdered carbon, permanganate, or, in extreme cases, modification of treatment process train (Bernhardt, 1984)), would not be required below a certain concentration range. In addition, costs of implementation, once required, would not be strongly dependent upon algal concentration, especially if substantial capital investments are involved.

Based upon the South African study, the frequency of severe nuisance conditions, as defined by chlorophyll *a* concentrations exceeding 30 ppb, is used as a measure of use impairment. For simplicity, the frequency of severe nuisance conditions is termed "bloom frequency", although these values do not necessarily correspond to biological definitions of the term.

FREQUENCY DISTRIBUTION MODELS

Frequency distribution models can be used to predict bloom frequency, given certain parameters. Chlorophyll *a* distributions tend to be skewed towards high values. The standard deviation increases in rough proportion to the mean (Walker, 1983b; Knowlton et al. 1984; Walmsley, 1984), which is typical of variables which are log-normally distributed (Snedecor and Cochran, 1972). A two-parameter, log-normal distribution model is presented to demonstrate the methodology. Future refinements could consider testing alternative distribution functions. In some cases, chlorophyll *a* distributions tend to be somewhat more skewed than predicted by the log-normal model.

To relate bloom frequencies to mean concentrations, a measure of spread (standard deviation) is required. Three models are summarized for estimating the standard deviation as a function of the mean (Table 2). The first model is based upon a regression analysis of data from Vermont lakes (Fig. 2). The regression indicates that standard deviation is proportional to the 1.21st power of the mean. The second model is based upon a log-scale variance component analysis of data from Corps of Engineer reservoirs (Walker, 1980, 1981). The third model is based upon a linear-scale regression analysis of data from South African reservoirs (Walmsley, 1984). Despite the fact that the three models have very different data sets, averaging periods, and functional forms, they give similar predictions of bloom frequency at a given mean chlorophyll *a*.

Table 2.—Statistical models for chlorophyll *a* variability.

	Model Number 1
References:	Walker, 1983b, 1984
Water Bodies:	Vermont Lakes
Data Set:	148 Station-Years
Sampling Freq.:	Weekly
Averaging Period:	June-August
Model:	$SA = .29\ MA^{1.21}$ Based upon log-scale regression of SA on MA ($r^2 = .84$, $SE^2 = .026$, $\log_{10}$ scales)
	Model Number 2
References:	Walker, 1980, 1981
Water Bodies:	U.S. Army Corps of Engineer Reservoirs (Nationwide)
Data Set:	258 Station-Years
Sampling Freq.:	3–4 Samples/station-year
Averaging Period:	April–October
Model:	$SL = .62$ Based upon among date variance component of reservoir-mean chlorophyll *a* values, log scales
	Model Number 3
Reference:	Walmsley, 1984
Water Bodies	South African Reservoirs
Data Set:	34 Reservoir-Years
Sampling Freq.:	Weekly, Biweekly, Monthly
Averaging Period:	annual
Model:	$SA = .95\ MA - 1.68$ Based upon regression of SA on MA, linear scales ($r^2 = .85$)

Symbols Defined in Table 3

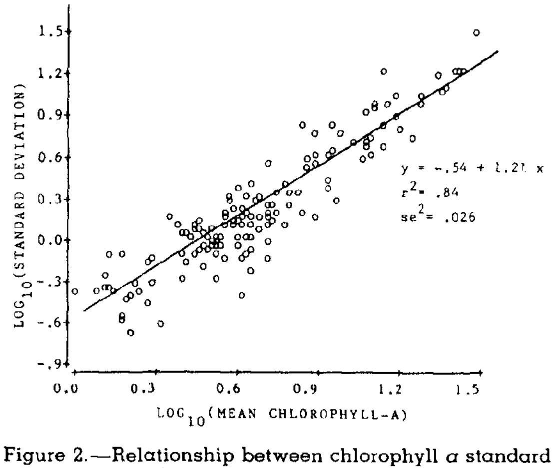

Figure 2.—Relationship between chlorophyll *a* standard deviation and mean for Vermont lakes (Units ppb).

Table 3.—Algorithm for calculating chlorophyll *a* frequency distributions.

SYMBOLS:

MA,SA = arithmetic mean and standard deviation of chlorophyll *a* (mg/m^3)
ML,SL = mean and standard deviation of $\log_e$ (chlorophyll *a*, mg/m^3)
Z = standard normal deviate (mean = 0, standard deviation = 1)
F(Z) = integral under standard normal curve from Z to infinity
C = instantaneous chlorophyll *a* value (mg/m^3)
V,W,X = variables used in calculating cumulative distribution function

ALGORITHM:

For a log-normal distribution, the following equations estimate arithmetic moments (MA,SA) from log-scale moments (ML,SL) (Aitchison and Brown, 1963):

$MA = \exp(ML + .5\ SL^2)$
$SA^2 = MA^2\,[\exp(SL^2) - 1]$

or vice-versa:

$ML = \log_e(MA) - .5\ SL^2$
$SL^2 = \log_e[1 + (SA/MA)^2]$

The percent of the chlorophyll *a* distribution exceeding a given chlorophyll *a* criterion (C*) can be calculated from:

$Z = [\log_e(C^*) - ML]/SL$
$\mathrm{Prob}\,(C > C^*) = F(Z) \times 100\%$ = percent of samples exceeding C*

F(Z) can be derived from statistical tables (Snedecor and Cochran, 1972), or estimated from the following empirical equation for the normal distribution:

$V = \exp(-Z^2/2)/2.507$
$W = (1 + .33267\,|Z|)^{-1}$
$X = V\,(.4361684\ W - .1201676\ W^2 + .937298\ W^3)$
If $(Z > 0)$ then: $F(Z) = X$
else: $F(Z) = 1 - X$

The Vermont and South African models are based upon arithmetic moments (mean and standard deviation). Transformation functions can be used to estimate the log-scale moments from the linear-scale moments (Aitchison and Brown, 1963) (Table 3). Using an algorithm for calculating extreme value frequencies (Table 3), it is possible to estimate the percent of the time nuisance levels are experienced, as a function of the arithmetic mean (or geometric mean) chlorophyll *a* for any of the empirical variance models (Table 2) and for any definition of nuisance level. If more complicated use impact models are employed (such as, Model 2 in Fig. 1), the calculations become somewhat more involved. They would entail numeric integration of the product of the frequency distribution function and the impairment index over the entire range of chlorophyll *a* concentration.

RESULTS

Observed and predicted bloom frequencies are plotted as a function of mean chlorophyll *a* (Fig. 3). These plots are based upon the Vermont model and data set, the most uniform and intensive of the three data sets described in Table 2. Three alternative definitions of bloom frequency are shown (20, 30, and 40 ppb)(Fig. 3). These plots demonstrate the ability of the methodology to predict extreme value frequencies as a function of the mean. Agreement is generally better for the 20 and 30 ppb criteria. The model tends to under-predict the frequency of concentrations exceeding 40 ppb, but is not extensively tested because chlorophyll *a* concentrations exceeding 40 ppb were detected in only 9 of 148 station-years tested.

Predicted bloom frequencies are plotted as a function of the arithmetic mean chlorophyll *a* concentration (Fig. 4). Results are shown for each of the three variance models in Table 2 and for three definitions of nuisance level (20, 30, 40 ppb). Predicted frequencies are not very sensitive to the particular variance model employed.

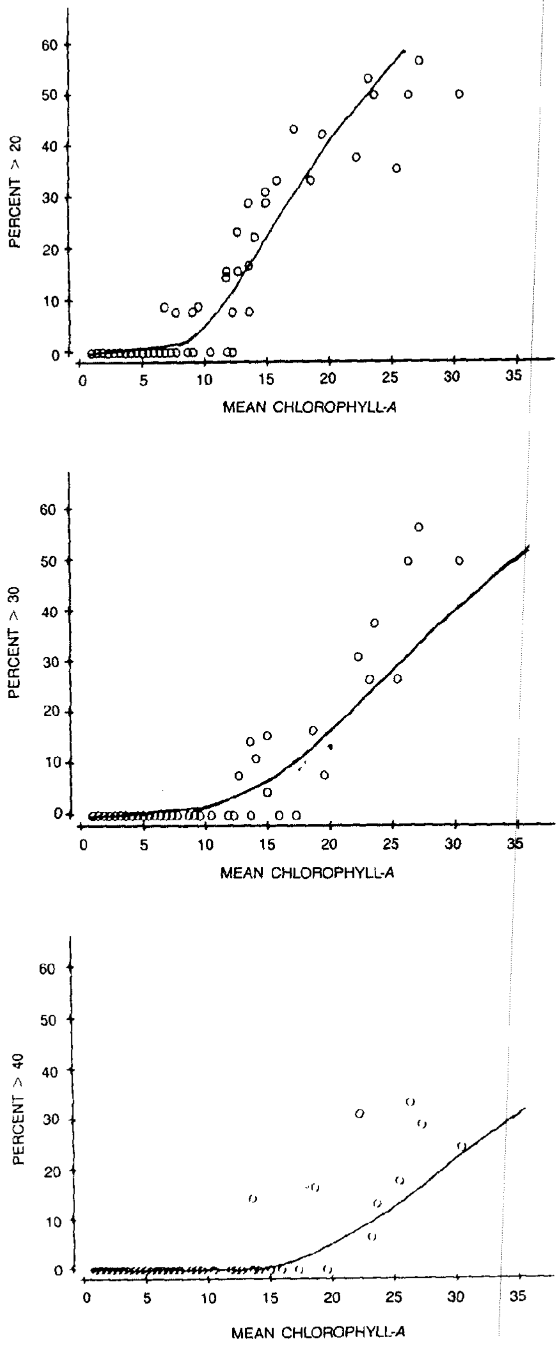

Figure 3.—Observed (o) and predicted (−) extreme value frequencies as function of mean chlorophyll *a* (ppb) for Vermont lakes.

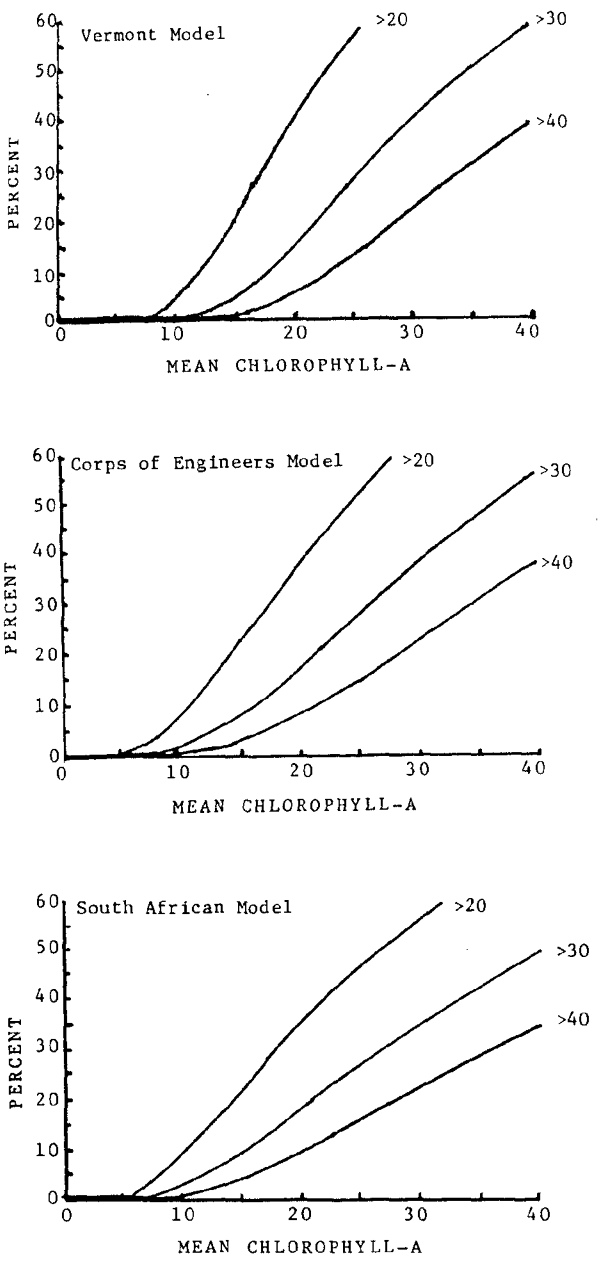

Figure 4.—Predicted extreme value frequencies for three variance models and three definitions of nuisance conditions.

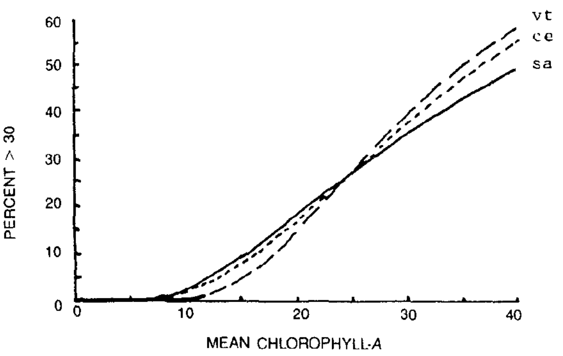

Figure 5.—Predicted frequencies of chlorophyll *a* concentrations exceeding 30 ppb as a function of mean chlorophyll *a* for three variance models (vt = Vermont model, ce = Corps of Engineers model, sa = South African model).

For Walmsley's definition of severe nuisance condition (30 ppb), the use impact is approximately a linear function of mean chlorophyll *a* and has an x-intercept of approximately 10 ppb (Fig. 5). Results can be summarized by the following equation for mean chlorophyll *a* concentrations between 10 and 40 ppb:

$$\text{Percent} > 30 = 1.83\ (\text{Mean Chl. } a{-}10)$$

For mean concentrations less than 10 ppb, expected bloom frequencies and resulting use impairment are minimal for a system with average variability. The statistically derived intercept agrees with subjective definitions of the mesoeutrophic boundary (Table 1).

As expected, the intercept is somewhat lower for the 20 ppb criterion and higher for the 40 ppb criterion. Qualitatively, however, the curves are of similar shape. The intercept represents a logical focus for developing lake water quality criteria. Systems with "average" variability at a given mean concentration are illustrated in Figures 4 and 5. Additional analysis is required to incorporate the effects of divergence from the predicted standard deviations in Figure 2.

CONCLUSIONS

If the impacts of eutrophication on water uses can be expressed in terms of the frequency of nuisance-level concentrations or blooms, use impairment in the lake with a mean chlorophyll *a* of 20 ppb would be much more than twice that in a lake with 10 ppb, which, in turn, would be little different from a lake with 5 ppb. This nonlinear relationship is not immediately obvious in classifying lakes and reservoirs based upon mean chlorophyll *a* or trophic state indices. While, from a biological point of view, it may be reasonable to consider the distribution of mean chlorophyll *a* values across lakes as a "continuum", this concept may be somewhat misleading for considering use impairment because of the intercepts in Figures 4 and 5.

If fisheries' benefits can be demonstrated from increasing chlorophyll *a* concentrations, this analysis suggests that these benefits will not be in conflict with protecting other water uses, provided that mean chlorophyll *a* values remain below the use impairment intercept.

The framework demonstrated here permits estimation of extreme value frequencies from existing compilations of mean chlorophyll *a* data. Site-specific data can be used to develop user response models and frequency distribution models appropriate for a given region and water uses. The potential for nonlinear responses of use impairment to increasing mean chlorophyll *a* should be considered in adopting lake-management goals.

REFERENCES

Aitchison, J., and J.A.C. Brown. 1963. The Lognormal Distribution. Cambridge Univ. Press.

Bernhardt, H. 1984. Treatment disturbances with water out of eutrophic reservoirs as a consequence of extensive algal development. Pres. at IWSA Conference, Monastir/Tunis, 1984.

Dobson, H.F.H., M. Gilbertson, and P.G. Sly. 1974. A summary and comparison of nutrients and related water quality in Lakes Erie, Ontario, Huron, and Superior, J. Fish. Res. Board Can., 31: 731–38.

Knowlton, M.F., M.V. Hoyer, and J.R. Jones. 1984. Sources of variability in phosphorus and chlorophyll and their effects on use of lake survey data, Water Resour. Bull. 20 (3): 397–408.

National Academy of Science. 1972. Water Quality Criteria, Washington, D.C.

Organization for Economic Cooperation and Development. 1982. Eutrophication of waters—monitoring, assessment, and control. Synth. Rep. from the OECD Coop. Programme Eutroph., OECD Publications, Washington, D.C.

Snedecor, G.W., and W.G. Cochran. 1972. Statistical Methods, 2nd ed. Iowa State Univ. Press, Ames.

U.S. Environmental Protection Agency. 1974. The Relationships of Phosphorus and nitrogen of Northeast and North-Central Lakes and Reservoirs. Natl. Eutroph. Surv. Work. Pap. No. 23.

Wagner, K.J. 1984. Incompatibility of lake management objectives. Lake Line. 4(2): 7. N. Am. Lake Manage. Soc., Washington, D.C.

Walker, W.W. 1980. Analysis of Water Quality Variations in Reservoirs: Implications for Monitoring and Modeling Efforts. *In* H.G. Stefan, ed. Surface Impoundments. Am. Soc. Civil Eng., New York.

———. 1981. Empirical Methods for Predicting Eutrophication in Impoundments, Rep. 1, Phase I: Data Base Development. TR E-81-9, U.S. Army Corps Eng., Vicksburg, Miss.

———. 1983a. Significance of eutrophication in water-supply reservoirs. J. Am. Water Works Ass. 75 (1): 38–42.

———. 1983b. Variability of Trophic State Indicators in Vermont Lakes and Implications for LEAP Error Analyses. Prepared for Vt. Agency Environ. Conserv., Water Qual. Div., Montpelier.

———. 1984. Eutrophication factors and objectives for Cherry Creek Reservoir. Prep. for City of Greenwood Village, Colo. in support of testimony before Colo. Water Qual. Contr. Comm. regarding phosphorus standard for Cherry Creek Reservoir.

Walmsley, R.D. 1984. A chlorophyll *a* trophic status classification system for South African Impoundments. J. Environ. Qual. 13 (1): 97–104.

Walmsley, R.D. and M. Butty. 1979. Eutrophication of rivers and dams. VI. An investigation of chlorophyll–nutrient relationships for 21 South African Impoundments, Contr. Rep., Water Res. Comm., Pretoria, South Africa.