Wisconsin Lakes Study (II): Statistical Covariance Analysis of Physical and Chemical Data $\sim 10^{-10}$ km s $^{-1}$

 \sim \sim

 \sim \sim

 \sim \sim

 $\sim 10^{-11}$

 ~ 100

 \mathcal{R}

T ^

 \sim

Robert Clifford, Jr., John Wilkinson, Nicholas L. Clesceri, Keith A. Booman and Richard I. Sedlak, Jr.

 \spadesuit

January 1985

 $\mathcal{O}(\mathcal{O}_\mathcal{O})$. The $\mathcal{O}(\mathcal{O}_\mathcal{O})$

 α

ABSTRACT

Statistical covariance analysis of data collected by the Soap and Detergent Association during its Wisconsin Lakes Study assessed the effectiveness of an interim detergent phosphorus ban. The covariance of basic limnological parameters between lakes expected to be impacted by a detergent phosphorus ban and lakes not impacted by the ban was examined. A statistically significant change in either total phosphorus or chlorophyll-a concentrations associated with the imposition of the ban was not observed on any of the lakes studied. An increase in secchi depth temporally associated with the ban was noted on two of the study lakes. For these two lakes, however, this change is unlikely to have resulted from imposition of the ban since statistically significant changes in chlorophyll-a were not observed. The covariate analysis also quantified the degree to which temporal fluctuations in total phosphorus, chlorophyll-a and secchi depth measurements may be observed between geographically proximate lakes.

INTRODUCTION

A previous paper (Clifford, et. al., 1985) has described the Wisconsin Lakes study conducted by the Soap and Detergent Association. Six lakes which were expected to receive a significant percentage of their phosphorus loadings from wastewater treatment plant effluent, along with one lake expected to receive effluent from domestic wastewater treatment systems, were monitored from 1978 through 1983. These lakes were considered as "test" lakes in guaging the effectiveness of the detergent phosphorus ban. \ln addition, monitoring was conducted upon three lakes not impacted by wastewater effluent. These lakes represented "reference" points from which water quality data would be unaffected by intervention of the ban. Lakes were grouped into test-reference pairs based upon geographic proximity.

Temporal plots of secchi depth, chlorophyll-a concentration, and total phosphorus concentration exhibited, in some cases, an observable degree of "tracking", or synchronic variation of the data, between test and reference lakes. It was expected that this "tracking" would be observed, in part, because seasonal or climatic factors which cause day to day fluctuations in water clarity and chemical makeup of the water column uniformly impact small geographic areas. Evidence of a monotonic shift in water quality parameters associated with imposition of the detergent phosphorus ban was not discernable, however, from these temporal plots of the data or from yearly means. A means was sought by which the ban, imposed on a trial basis, might be evaluated for any impact it had upon the seven test lakes. Information collected by the monitoring program was therefore input to a computer data base and statistical procedures were employed to detect water quality trends on the test lakes relative to the reference lakes.

Typically, trend analysis of water quality data is hampered by several factors, among them missing values, values below detection limit, seasonality, and the non-normality of parameter distributions (Hirsch, Slack and Smith, 1982, Van Belle and Hughes, 1984). As a result, non-parametric statistical methods are usually employed to determine time-related variations in water quality. It has also been reported that an extensive data record is necessary in the assessment of lake restoration programs in order to in statistical power level if parametric tests are used (Trautman, McCulloch and Oglesby, 1982). These studies, however, assume that a monitoring record is available only for lakes impacted by a phosphorus control method. The use of a test-reference lake approach establishes a baseline which characterizes the seasonality of water quality data collected from a geographic region and permits the identification of trends that might otherwise go undetected.

Parametric procedures may then be employed if reasonably continuous data exhibit distributional normality.

ESTABLISHMENT OF TEST VARIABLES

During this study the following water quality parameters were monitored: temperature and dissolved oxygen, secchi depth (SD), alkalinity and pH, and the concentrations of chlorophyll-a (CHLA), total phosphorus (TP), total oxidized nitrogen (TON), ammonia (NH4), and total kjeldahl nitrogen (TKN). In addition, measurements were made at two or more sites on each of the lakes. From this extensive data set a subset was sought which would be applicable to parametric statistical analysis and usefull in determining the detergent phosphorus ban's impact. In selecting this subset a decision had to be made as to which of the parameters monitored would best be indicative of the lake's water quality and from which of the multiple sites could measurements be accepted as characteristic of a particular lake. The concentrations of total phosphorus and chlorophyll-a as well as secchi depth have wide acceptance as standard water quality indicators and were therefore initially considered as test variables. Each variable was considered separately in the statistical analysis.

In order to determine if there was significant spatial variability in measurements made from separate sites on a single take the standard Pearson Product-Moment Correlation Coefficent (r) was calculated for log values of the monitored parameters between sites. Correlations were made for each lake separately as well as for the whole data set, and the results appear in Table 1. Samples and measurements were taken within the epilimnion at a location determined from bathymetric charts as the deepest point on the lake, as well as at one or two other locations believed to represent possible hydrologic sub-basins. In most, but not all cases, the deepest point of the lake was also

located approximately at the center. The best correlations between sites were observed for secchi depth (SD) and total kjeldahl nitrogen (TKN) concentration. This probably occurred because these measurements have the least sensitive detection limits (secchi depth was recorded to the nearest foot and TKN was recorded at the ppm level; all other chemical analyses were 2 recorded at the ppb level). The value of r^2 may be interpreted as the amount of variation observed at one site that is explainable by its relation with the other site. It may be seen that for secchi depth, total phosphorus (TP), and chlorophyll-a (CHLA) the value of r is always greater than or equal to 0.74, consequently at least 55% (r^2 = 0.55) of the variability ir t measurements made at one site is coincidentally observed at, and explained by, its relation to any other site. These correlations strongly indicate that the data recorded at the deepest site is representative of data recorded at any other site and would be acceptable for use in the statistical comparisons to be made between lakes. Incorporation of the data taken from additional sites into the statistical analysis would have had to have been done on a selective basis, since a variable number of additional sites was sampled on each lake. Also, it would be impossible to determine if this data would more uniquely characterize each lake or simply add the variability incurred during chemical analysis into the data base.

Therefore it was determined that the subset of the Wisconsin lakes monitoring data to be used in statistical analyses would be the concentrations of total phosphorus, chlorophytl-o and the secchi depth recorded from the epilimnion at a location determined to be the deepest point of the lake.

SUITABILITY OF THE DATA BANK

All data analyses were completed using an IBM 3081D main frame computer under the Michigan Terminal System (MTS). The statistical package

MIDAS, the Michigan Interactive Data Analysis System (Statistical Research Laboratory, The University of Michigan, 1976) was used in determining variable normality. All regressions were run using BMDP statistical software (Dixon, et. at., ed., 1983).

Testing for Variable Normality

n order to justify the use of parametric statistical procedures in the trend analysis of this monitoring record, each of the parameters was tested for distributional normality. An example of the distributions found is presented in Figure 1. Similar to the distribution found for total phosphorus (Figure 1(a)), natural distributions for secchi depth and all of the chemical species monitored were skewed below the sample mean. Application of the log function to the data, and subsequent standardization, produced distributions similar to that shown in Figure 1(b), which more closely resembles the "bell-shape" of a Gaussian normal distribution. For phosphorus, 74.9% of the values recorded fell between +1 standard deviation and 92.0% of the values fell between +2 standard deviations, confirming that the log distribution is normal. Log normal distributions were determined for secchi depth, chlorophyll-a concentration, the concentrations of all of the nitrogen species measured, alkatinity and pH. This suggests that the use of parametric statistical tests based upon normality assumptions wilt be reasonable when these variables are considered.

Completeness of the Data Set

The data set collected during this study is relatively complete. Only 36 measurements out of 3990, or less than one-tenth of one percent, were recorded as missing or below the detection limit. This is due to the care taken in sample handling and analysis and to the fact that most of the chemical analyses were conducted at the low ppb level of detection.

STATISTICAL METHODOLOGY

Covariance Analysis for Each Test Lake

For each test-reference lake pair, a covariance analysis was performed upon the values recorded for secchi depth, total phosphorus and chlorophyll-a concentration using the following model:

$$
\log y_{ti} = \beta_0 + \beta_1 \log y_{ri} + \beta_2 B_i + E_i
$$
 (1)

where:

 y_{ti} represents the i-th observation on the test lake

 y_{ri} represents the corresponding i-th observation on the reference lake

 B_i is a dummy variable with value 0 for pre-ban observations and value 1 for post-ban observations. Table 2 lists the salient features of the analysis. The quantity of interest is β_2 , which represents the post-ban shift in the test lake response after the modelled relationship with the reference lake has been considered. Only for Elk Lake and Townline Lake, and for both cases only for secchi depth, was this shift statistically significant.

The strength of the linear relationship between logarithmic values of water quality parameters measured on test and reference lake pairs is indicated by the significance level of the slope coefficent, β_1 , and the value of r^2 , which may be interpreted as the proportion of variability the relation between the test and reference lake. For 14 of the 18 cases the linear relationship is significant at approximately the five percent level,

explaining between 13% to 54% of the variability observed in water quality measurements made between geographically proximate test-reference lake » pairs .

Combined Covariance Analysis for All Test Lakes

For a simultaneous examination of all test lakes, a covariance analysis was performed using the model:

$$
\log y_{ti} = \beta_0 + \beta_1 \log y_{ri} + \frac{6}{\beta_1} \alpha_j D_{ji} + \frac{6}{\beta_1} \alpha_j D_{ji} \log y_{ri} + \delta B_i + \frac{6}{\beta_1} \alpha_j D_{ji} \log y_{ri}
$$
\n
$$
\frac{6}{\beta_1} f_j D_i B_i + E_i
$$
\n(2)

where:

 $B_i = 0$ for pre-ban or 1 for post-ban observations

 $D_{\small \boldsymbol{\mathsf{j}}\, \boldsymbol{\mathsf{i}}}$ is coded either 1 or 0 as f

The log test lake response can be considered as the sum of ^a general intercept, a linear component relationship with the log reference lake response, an adjustment in the intercept for the specific test lake, an adjustment in the slope for the specific test lake, an adjustment in the general intercept for the pre/post-ban period, and an adjustment in the intercept, for a specific test lake, for the pre/post-ban period. Another way of expressing this is that one first adjusts the total test lake response variability for a potential relationship with the corresponding reference take response and for individual test lake differences and then examines for the effect of imposition of the ban. In this analysis additional degrees of freedom become available for statistical testing purposes, giving a greater capability of detecting differences. The results are presented in Table 3. Secchi depth showed a detectable variation between the pre- and post-ban values at the five percent level of significance. However, only an additional 7% of the variability observed between test and reference lakes in this model was explained by the pre/post-ban intercept adjustment for secchi depth. Overall, the model accounted for more of the variability for secchi depth (71%) and total phosphorus (51%) than for chlorophyll-a (31%).

Estimating Actual Water Quality Fluctuations

In order to relate the observed changes in the regression coefficents in terms of the parameters being modelled the regression equations for secchi depth on test lakes Elk and Townline were used to evaluate the expected test lake secchi depth, at the mean reference lake secchi depth, for both the preand post-ban periods. At this reference lake level, the pre/post ban change in secchi depth for Elk Lake was calculated to be an increase of 0.35 meters, while an increase of 0.64 meters was calculated for Townline Lake. It should be noted that the model assigned statistical significance to these small changes

in water clarity at the five percent level.

Relationship Between Trophic Status and Lake "Tracking"

Table 4 ranks the test lakes according to their relationship with their corresponding reference lake in terms of the values of r^2 , which considered as the fraction of the variability in the test lake response which can be explained by a linear relationship with the reference lake variable. Adjacent to this column is a ranking of the lake pairs in terms of the value of β_0 , the value of the model intercept coefficent, which may be considered indicative of the mean ratio between the two lakes for a particular test variable. The ranking has been done in terms of the values for total phosphorus, and it may be seen that the lake pairs are ordered in the same 2 way for both r^{\pm} and \upbeta_{0} . In other words the goodness of fit of a li relating temporal changes in total phosphorus concentration between two takes located in the same geographical proximity is inversely related to the mean ratio of their concentrations. Such a clear cut relation was not observed, however, for either secchi depth or chlorophyll-a measurements.

Re-evaluation of Nutrient Loadings to the Study Lakes

An initial criteria used in selecting test lakes was the condition that approximately 30% or more of the total phosphorus budget to the lake be attributable to treated wastewater effluent. In order to quantify the expected benefits of the detergent phosphorus ban, nutrient budgets to the test lakes were recalculated as a part of this study (Clesceri, et. al., 1985). The portion of these revised budgets assigned to wastewater treatment plant effluents is presented in Table 5 and show that an affluent loading impacts only four lakes. Swan, Elk, Butternut and Townline. At the onset of the study expected wastewater loadings to the other test lakes fulfilled the 30% criteria. The change was due, in the case of Townline Lake, to the

upgrading of its treatment plant from primary to secondary treatment and, in the cases of Balsam, Moss and Enterprise lakes, to the reassessment of expected loadings from land disposal systems. In light of these loading projections it might be assumed that an observable effect of the detergent phosphorus ban would not be seen on Enterprise, Moss and Balsam lakes and the covariate statistical analysis confirms this assumption. However, it would seem unlikely that the detergent phosphorus ban, which affected only 3-4% of the wastewater treatment plant phosphorus loading to Townline Lake, was the causative factor involved in the increase in secchi depth noted in the covariate analysis.

Analysis of Data Collected Following the Recission of the Ban

Monitoring was continued on three lakes in 1983, following recission of the detergent phosphorus ban. However, of these three, only one test-reference lake pair, Butternut Lake and Teal lake, was included. A covariance analysis was conducted in a manner similar to the individual lake analysis, categorizing pre-ban, during-ban, and post-ban data. A trend in . either secchi depth, total phosphorus or chlorophyll-a data temporally related to recission of the ban was not noted for the Butternut-Teal lake pair.

DISCUSSION

For all cases of lake pairs and parameters, the only significant change in water quality parameters temporally related to the interim detergent phosphorus ban was an increase in secchi depth for Elk Lake, relative to its reference lake Teal, and for Townline Lake, relative to its reference lake Little Bearskin. It should be noted that a corresponding statistically significant change in total phosphorus or chlorophyll-a concentration was not observed in either case, and so associating the increase in water clarity with either nutrient levels or algal concentrations is inappropriate. In addition,

the linear relationship between Townline Lake and its reference, Little Bearskin Lake, is not statistically significant for secchi depth and so the indication that there was a relative change between the two related to the imposition of the ban is inconclusive.

The combined covariance analysis also indicates that secchi depth was the only parameter found to change after the ban was imposed. However, the degree to which this relation affects the model (a 7% increase in it's predictive capacity) and the lack of a coincident flux in total phosphorus or chlorophyll-a measurements over the same period of time leads to the conclusion that the ban was probably not the factor primarily involved in the change.

Of note from the individual lake analysis was the extent to which a simple linear relation attained statistical significance when secchi depth, total phosphorus and chlorophytl-a were considered between lakes in geographic proximity. Natural fluctuations in water quality are apparently so closely related to climatic and seasonal conditions that their fluxes may be reflected uniformly in different water bodies. Table 4 underlines the fact that the "goodness-of-fit" of this linear relationship is closely allied with the trophic differences between the two lakes. Analogously, one might consider the degree to which settling dust is more apparent on a clean dinner plate than on a plate already soiled. It would seem reasonable that small changes in phosphorus loadings would have a more immediate and visible affect upon lakes tending towards oligotrophy. A comparison of Table 4 with Table 2 shows also that, for lake pairs, the slope coefficent of the logarithmic linear relationship for a time series of coincident measurements approaches unity as the takes approach a similar overall mean total phosphorus concentration. The conclusion drawn is that the test-reference lake approach to trend analysis

^

will prove most successfull when the paired lakes are initially of a similar trophic state, as reflected in mean values of the parameters being considered.

The combined analysis points out the relative sensitivity of the three measurements, total phosphorus, chlorophyll-a and secchi depth, as predictive tools is assessing water quality. It may be seen from Table 3 that the least variability among measurements (as exhibited in values for the sum of squares) occurs for secchi depth, followed by total phosphorus and chlorophyll-a. This would seem reasonable since the only directly measured parameter is secchi depth, a simple visual observation usually quantified to no more than the nearest foot. The assessment of total phosphorus is an indirect determination by which the absorbance of a color complex is related, via Beer's Law, to concentration (Murphy and Riley, 1962). The determination of chlorophyll-a in the water column is also an indirect measure, relating the absorbance of pigment remaining after a rigorous extraction technique to concentration on a linear scale (APHA, 1980). Given also that the analysis of total phosphorus and chlorophyll-a is carried out to the ppb level, the variability observed in the sums of squares for each of the parameters is ikely to be more a function of the analytical methodology than fluxes naturally occuring in the water column. It is interesting to note that the most ^t stable indicator is also the most classical, as secchi depth records are typically the only measurements which date back a significant number of years from the present.

The success of this study was closely aligned with the proper selection of lakes to monitor. This process was conducted, prior to 1978, with the best available information on lakes which were expected to receive a significant percentage of their phosphorus budget from wastewater treatment plant (WWTP) effluent and therefore be impacted by detergent phosphorus.

Recalculated nutrient budgets have shown that only one lake (Swan) can credit more than 30% of its phosphorus loading to WWTP effluent and, in fact, four of the seven "test" lakes would virtually not be impacted at all by a ban on detergent phosphorus. In most respects the covariance analyses described here substantiate the expected abscence of a trend in water quality data temporally related to imposition of the ban for these lakes. In the case of Townline Lake, where a small relative increase in secchi depth was temporally related to the ban period, the recalculated WWTP loading of 8% phosphorus (or 3-4% detergent phosphorus) undermines any attempt to form a cause and effect link between increasing water clarity and the Wisconsin detergent phosphorus ban. Likewise, simlarly relating the relative increase in secchi depth determined by the model for Elk Lake would also fall under suspicion, although the extent to which the ban was expected to reduce phosphorus loadings to Elk Lake was much greater (a 7-11% decrease in the total budget). Because corresponding decreases in total phosphorus or chlorophyll-a . concentrations were not coincidentally observed upon either lake other factors affecting water clarity may have caused the increase in secchi depth. Lambou, et. al. (1982) have noted the effect the prescence of dissolved color and/or suspended particles have upon clarity of the water column. Also, Elk Lake was found to have a short hydraulic residence time (less than 5 days). Essentially, this would tend to reduce the importance of ambient nutrient concentrations in determining algal production and, hence, water clarity, because of an increasing probability that rapid washout of nutrient and solids laden runoff more often than not had a determining effect upon clarity in the water column.

Overall, this study has jointly determined both the impact of a detergent phosphorus ban upon a set of lakes as well as whether or not an impact would have been expected. Determining the latter in advance of initiating a costly monitoring survey is the focus of a related paper (Clesceri, et. al., 1985a). However, it has been seen that there was no improvement in water quality, reflected in lower total phosphorus or chlorophyll-a concentrations or increased secchi depth, upon the ten lakes selected for this study as a result of an interim detergent phosphorus ban.

CONCLUSION

An examination of the data set collected during the Wisconsin lakes monitoring study determined that a usefull subset for statistical analysis be the values for secchi depth, total phosphorus and chlorophyll-a concentration recorded at the location of the deepest point on the lakes. The data set was complete/ and a log normal distribution was confirmed for the parameters included in the analysis, justifying the use of a parametric statistical test. Covariance analysis, utilizing a linear regression model for each test-reference lake pair, determined that at the five percent level of significance only an increase in secchi depth for two of the lakes. Elk and Townline, was found to have occured coincident with the term of a detergent phosphorus ban. Since improvements in total phosphorus and chlorophyll-a concentration did not occur, the observed improvement in water clarity in these two lakes does not appear to be related to the imposition of the ban. In addition, the ban affected only 3-4% of the total phosphorus loading to Townline Lake, diminishing its significance as a causative factor in the observed increase. Also, a statistically significant linear relationship was determined for at least one of the parameters secchi depth, chlorophyll-a or total phosphorus concentration between each of the seven lake pairs.

Literature Cited

- Clifford, Jr., R., M.A. Meyer, R.I. Sedlak and N.L. Clesceri. 1985. 'Wisconsin Lakes Study (I): A Multiple Lake Method to Analyze Temporal Variations'. In press, this journal.
- Hirsch, R.M., J.R. Slack and R.A. Smith. 1982. 'Techniques of Trend Analysis for Monthly Water Quality Data'. Water Resources Research,18(1): 107-121.
- Van Belle, G. and J.P. Hughes. 1984. 'Nonparametric Tests for Trend in Water Quality'. Water Resources Research, 20(1): 127-136.
- Trautmann, N.M., C.E. McCulloch and R.T. Oglesby. 1982. 'Statistical Determination of Data Requirements for Assessment of Lake Restoration Programs'. Canadian Journal of Fisheries and Aquatic Sciences, 39: 607-610.
- Clesceri, N.L., S.J. Curran and R.l. Sedlack. 1985. 'Nutrient Loads to Wisconsin Lakes, Part I: Nitrogen and Phosphorus Export Coefficents'. In press.
- The Statistical Research Laboratory, The University of Michigan. 1976. Elementary Statistics Using MIDAS. The Statistical Research Laboratory. The University of Michigan. 300 pp.
- Dixon, W.J., ed. 1983. BMDP Statistical Software. 1983 Printing with Additions. University of California Press. Los Angeles. 733 pp.
- Murphy, J. and J.P. Riley. 1962. 'A Modified Single Solution Method for the Determination of Phosphate in Natural Waters'. Analytica Chimica Acta, 27:31-36.
- American Public Health Association. 1980. Standard Methods for the Examination of Water and Wastewater. APHA-AWWA-WPCF. 1134 pp.
- Clesceri, N. L., S.J. Curran and R.l. Sedlack. 1985a. 'Nutrient Loads to Wisconsin Lakes, Part II: Relative Importance of Nutrient Sources'. In press.

TABLE 1: Pearson Product-moment Correlation Coefficients (r) Between Sampling Sites for Log Values of the
Physical/Chemical Parameters Monitored

Figure 1(b). Distribution Histogram, Standardized log TP Data

^

Table 2: Covarlance Analysis for Individual Test Lakes

.

-+<. .^*- ? * »

Significant at the $5%$ l

4t

VI TARA KONGRESI MENDELEN ANAR MENDELERAN DENGAN TAHUN 1999-YANG ANG AKTOR AT PANGAN AT MANGER AT MANGER AT MA

Table 3: Combined Covariate Analysis

Table 4: Tracking of a log-linear relationship between test and reference lakes. Percent variation (R-squared) explained by log-linear relationship compared with intercept value of the log-linear model (for total phosphorus data).

TOTAL PHOSPHORUS

Lake	Reference	r^2	$\mathbf{B}_{\mathbf{O}}$
Moss	Little Bearskin	.54	$-.28$
Enterprise	Ħ	.47	$-.41$
Townline	Ħ	.31	.85
Balsam	Teal	.31	.89
Butternut	Ħ	.24	1.29
Elk	Ħ	.13	1.36
Swan	Fish	\star	1.89

^{*}

Indicates that relationship is not significant at the five-percent level.

Table 5: Municipal WWTP Phosphorus Loadings to Lakes.

 $\sum_{\substack{\lambda,\lambda'\in\mathcal{P}^{\mathcal{P}_{\lambda},\lambda'\\ \lambda'\in\mathcal{P}_{\lambda'}\cup\mathcal{P}_{\lambda'}}}} \varphi_{\lambda,\lambda'}(x)$

Note: About 30% of wastewater phosphorus may b assumed to come from detergents

 ϵ