Title
A PATTERN ANALYSIS OF CLEAR LAKE PHYTOPLANKTON

Author
Sandusky, James C

Publication Date
1977-03-01

Submitted to
Limnology and Oceanography

LBL-6302
Preprint

Energy and Environment Division

A Pattern Analysis of Clear Lake Phytoplankton

James C. Sandusky and
Alexander J. Horne

March 1977

Lawrence Berkeley Laboratory University of California/Berkeley

Prepared for the U.S. Energy Research and Development Administration under Contract No. W-7405-ENG-48

DISCLAIMER

A PATTERN ANALYSIS OF CLEAR LAKE PHYTOPLANKTON

James C. Sandusky
Lawrence Berkeley Laboratory

and

Alexander J. Horne
Department of Sanitary & Hydraulic Engineering
University of California, Berkeley

This work was done with support from the U.S. Energy Research and Development Administration, Lake County Flood Control and Water Conservation District, the California Dept. Water Resources and the National Aeronautics & Space Administration at Ames, Cal., and a National Science Foundation Postdoctoral Fellowship grant.

ABSTRACT

In large lakes, it is usually impossible for logistic reasons to collect enough samples to analyze spatial patterns of algae, zooplankton, nutrients or physical factors using available methods. Yet such an analysis would be ecologically useful. Using results from synoptic measurements which most limnologists could gather, we had three objectives. First, to develop a statistical frequency distribution more suitable for aquatic plankton than existing ones. Second, to analyze the synoptic data using this distribution. Third, to use the results of our analysis to quantify phytoplankton spatial relations and to investigate physical and ecological factors which control them. We analyzed forty-three sets of data (about thirty-two samples per set) representing eighteen freshwater phtoplankton genera and obtained reliable results using three fitted parameters for each set. We explicitly separated local and regional variability, thus avoiding overestimates of regional variability (patchiness) caused by counting individual cells of colonial or locally associated species as independent units. Our patchiness estimates showed significant between-species variations which could be related to ecological and environmental factors.

INTRODUCTION

Spatial organizations of biotic communities are an important part of the dynamical structure of these communities. Floral and faunal terrestrial populations are often found in clumps or patches for a variety of reasons, including variations in nutrient supply (soil conditions) and water supply, limited dispersal in reproduction and protection afforded from predation. In the aquatic environment, these effects are blurred by advective circulation and turbulent mixing. Because mixing is incomplete, a good deal of spatial coherence exists. For marine and freshwater phytoplankton, patchiness is common and appears to be ecologically significant (Richerson, Armstrong and Goldman, 1970; Steele, 1974; Riley, 1976). A priori one would expect patchiness to be a disadvantage because of such drawbacks as nutrient depletion and predation by zooplankton or fungi. In spite of this, dense patches occur and one can speculate that patchiness may provide competitive advantages to some phytoplankton species. These might include shading of competing forms (Lund, 1965), generation and use of warm water surface lenses for rapid horizontal movement (Wrigley and Horne, 1975), secretion of algicides or chelation of toxic and beneficial metals, or avoidance of predation (Mullin and Brooks, 1976).

Spatial variations have been measured in a variety of ways. Early studies (Bainbridge, 1957 and references therein) were chiefly qualitative. Recent quantitative studies can be grouped into two categories according to sampling strategy: transects and synoptics. Sampling along

transects, has been used extensively to study spatial organizations of plankton (Margalef and Estrada, 1971). It is especially appropriate for systems with a linear gradient such as tidal zones and estuarine basins (Cassie, 1960), for systems with little transverse variation as is the case in many narrow freshwater lakes and reservoirs (Bittle, Grimm and Ochocki, 1965), or for marine systems where the area of study is large (Fasham, Angel and Roe, 1974).

In a synoptic study, water samples are collected from many sites representing a large area. This is a useful technique for studying systems where horizontal variation in more than one direction is present (Platt and Filion, 1975) or where it is desirable to understand circulation, topographic and inflow effects (Horne, Dillard, Fujita, Goldman 1972; Horne, Sandusky and Carmiggelt, 1977). Remote sensing by aerial photography (visible or infrared) shows promise as another way of studying a few aspects of two dimensional variability (Horne and Wrigley, 1975) particularly when used together with "water truth" measurements gathered from synoptics.

In this report we develop a systematic approach to interpret synoptic surveys using statistical methods to derive a distribution function for synoptic survey data. The resulting model is used to describe data on water chemistry and phytoplankton populations recorded during synoptic studies at Clear Lake, Calif. Parameters resulting from our analysis reveal interesting ecological features of these populations. We are

grateful to C.R. Goldman and P. Javornicky for providing us with some as yet unpublished survey data from 1969-70; to J. Neyman and P. Puri and E. Scott for contributions to the statistical theory. Useful discussions were also held with T. Powell and P. Richardson.

THEORY

An integral part of any quantitative experimental technique is the analytical methodology used to interpret the results. Sensible descriptive mechanisms can reduce experimental labor by eliciting maximal information from data and can provide insight into important ecological processes. In analyzing transect data, a major breakthrough came through the use of power spectral analysis to reduce data (Platt and Denman, 1975). Central to this application is the explicit use of between station distance as the variable which parameterizes population fluctuations. Use of power spectral analysis is also advantageous since many results from studies of hydrodynamic turbulence can be conveniently expressed in this framework and thus provide greater understanding of results.

For synoptic data, no such established methodology exists. Two dimensional power spectral analysis requires a far larger number of samples than are normally available in even the most ambitious synoptic survey. Reduction to one dimension using absolute distance between

stations as the spatial variable has obvious drawbacks. Advective current profiles in lake and estuarine systems are complicated. Because of topological irregularities and preferential wind direction, spatial fluctuations are seldom isotropic. Thus, when stations are not collinear, between station distance is not an accurate measure of "real" distance (in terms of water movement). Unless current patterns are well known, it is difficult, if not impossible, to say what this real distance should be.

Statistical models of spatial variations offer an alternate approach toward analysis of synoptic data. In terrestrial ecology, clustering models have been used extensively to understand two dimensional spatial fluctuations in floral and faunal populations (Patil, 1968; Peilou, 1969). These models have also been used in aquatic ecology to analyze transect sampling experiments (Barnes and Marshall, 1951; Cassie, 1963). In applying this method to analyze synoptic survey data, relative positions of stations are ignored. For most cases where current patterns are complicated or unknown, no loss of information occurs by this simplification.

Much statistical work on patchiness has been based on the Poisson distribution. Sampling a randomly dispersed population and tabulating frequencies of observed numbers of individuals results in a frequency distribution,

$$f(x|\mu) = e^{-\mu} \frac{\mu^x}{x!} \quad , \qquad (1)$$

where $f(x)$ = probability of finding x individuals and μ is the mean population density. A main feature of the Poisson distribution is equal variance and mean. For non-randomly dispersed populations, either mean exceeds variance (uniform) or variance exceeds mean (patchy). Fisher (1950) derived criterion for testing the significance of such deviations. Subsequent attempts to quantify these differences found several "patchiness indices" all based on the amount by which variance exceeds mean (Hill, 1973). Most can be related to the parameter,

$$k = \frac{(\text{mean})^2}{(\text{variance} - \text{mean})} \quad , \qquad (2)$$

which is the shape parameter for the negative binomial distribution.

In applying this work to aquatic ecology, two questions immediately arise: 1) which are the optimal parameters (indices) to use? 2) What is the most accurate way to estimate them? The answer to both questions lies in the correct choice of an underlying probability distribution. If, for example, population counts from a sampling experiment are well described by a negative binomial distribution, then the two parameters to use in describing the population are the mean and shape parameter (defined above) and the best way to estimate them is to fit the negative binomial distribution to the observed population frequencies. Thus, our procedure for analyzing experimental data is to first find an appropriate distribution function which contains descriptive parameters, then, to use the data to determine values for these parameters, and finally, to use these values to interpret the experiment. Choice of an optimal underlying distribution thus maximizes the information obtained from the experiment.

An important consideration in seeking a reliable probability distribution is that there is no unique one. Solutions to equations governing turbulent hydrodynamics and biological productivity do not exist. Even if they did, clustering is expected to manifest itself differently depending on system and type of species within a system. Thus, statistical distributions which worked well in describing terrestrial populations may not apply in the aquatic case. The procedure for constructing these models, however, is quite general and we will rely on it extensively here. In building a model, it is advantageous to utilize as much as possible, prior information, without which any model is merely an empirical device.

For the use of synoptic sampling, we begin by assuming the following. First, stations are indistinguishable in the sense that plankton are equally likely to be at any of them and whether or not they are, is a random process. Notable exceptions to this (e.g., diatoms which may be more likely to be at edge stations than center ones in lakes) are treated after the analysis is made. Second, stations are far enough apart so that water mass coherence is lost when water flows from one station to the next. These assumptions are equivalent to the assumption that population counts from each station form a sample space of independent, identically distributed events. To construct a distribution which describes these events, it is convenient to picture aquatic processes in terms of a length scale, s. Results from many studies (Kierstead and Slobodkin, 1953; Platt, 1972 and Powell et al., 1975) indicate the existence of three spatial regions:

(1) a clustering zone (10 m $\gtrsim$ s); (2) a turbulence zone (1000m $\gtrsim$ s $\gtrsim$ 10m); and (3) a patchiness zone (s $\gtrsim$ 1000m). The boundaries separating these regions are variable. Those for open ocean are expected to exceed those for estuaries or lakes. We assume stations are far enough apart for phytoplankton to show patchiness, otherwise we are observing effects primarily of water turbulence. At each station, populations may show evidence of local clustering. We therefore construct a model which describes clustering at small scales, and allow parameters to vary from station to station to describe patchiness. The intermediate zone is treated here, as a randomizing agent which assures the independence of station samples.

Small Scale Clustering

At a given station we are concerned with small scale variations in numbers. Work by Bernhard and Rampi (1965) and McAlice (1970) found significant overdispersion of phytoplankton on scales of 10 cm to 10 m. In part, this clustering occurs because the fundamental unit for phytoplankton is not always a cell but is sometimes an association of cells (chains, clusters, filaments, or grouping of filaments). In this case, one unit is observed infrequently (using the Poisson distribution as a standard) relative to zero units or many units. Clustering also results when algae are vertically stratified either passively by wind induced currents (Stavn, 1971) or actively through positive buoyancy control or swimming. Again, units are not spatially independent; plankton occur in spatial clusters of units.

We describe both cases by assuming that aggregates (a random number of individual units) are randomly distributed in space. Mathematically, this situation is conveniently described (Feller, 1950) by use of probability generating functions (pgf), defined by

$$G_X(t) \equiv E(t^X) \equiv \sum_{x=o}^{\infty} t^x P(x) \quad . \tag{3}$$

Here, G is the pgf for a probability distribution for the random variable, X. The probability of observing $X = x$ events is denoted by $P(x)$; $E(\cdot)$ is expectation. Probabilities may be found from the pgf by differentiating with the dummy variable, t:

$$P(x) = \frac{1}{x!} \frac{d^x}{dt^x} G_X(t)\Big|_{t=o} \quad . \tag{4}$$

For the Poisson distribution,

$$G_X(t|\mu) = e^{-\mu(1-t)} \quad , \tag{5}$$

where in this case X = number of clusters, and μ = cluster density.

If the pgf of the number of cells per cluster is denoted by $g_X(s|\lambda)$ with λ = average number of clusters, and clusters are distributed at random, it is easy to show (Feller, 1950) that the pgf for X = the number of individuals per sample is

$$G_X(s|\mu,\lambda) = e^{-\mu(1 - g_X(s|\lambda))} \tag{6}$$

There are many possible choices for the distribution of the number of cells per cluster. It is unlikley that one distribution will satisfy all cases. Certianly the case of fairly passive clus-

tering due to depth stratification as near a pycnocline is different from the case of clustering in multicellular forms. However, if we restrict ourselves to one free parameter, λ = number of cells per cluster, and in addition require: 1) a unimodal distribution, 2) variance proportional to $(\text{mean})^2$ for large means, 3) probability of a cluster with zero units is zero; then one distribution which satisfies these requirements is the goemetric distribution,

$$g_\chi(s|\lambda) = \frac{s}{\lambda - s(\lambda-1)} \quad . \tag{7}$$

The resulting distribution for number of individuals per sample is called the Polya-Aeppli distribution (Polya, 1931) and has been used to describe many diverse clustering processes. Other unimodal distributions are possible. For example, use of logarithmic or Poisson distributions for cluster sizes results in negative binomial or Neyman-A (Neyman, 1939), distributions for the number of individuals per sample

Large Scale Patchiness

We now consider what happens if we sample locations far enough apart to observe patchiness. The Poly-Aeppli distribution has two parameters, cluster density and cluster size and one or both of these will change. For simplicity, we will consider the phycologically more typical case where cluster density is a random variable and cluster size is fixed. Distributions where both parameters are random variables are interesting but tend to be mathematically intractable. If μ , the cluster density, is a random variable, we can make the following

requirements on the distribution function, $f(\mu)$: 1) $f(\mu)$ is unimodal and continuous over the open domain (o,∞) and 2) $f(\mu)$ depends on no more than two parameters. Two distributions which satisfy these requirements are log-normal and gamma. Both these distributions have found wide use in many diverse situations.

Here, we use the gamma distribution as it results in simpler formulaes, and, in the analysis of data it gave a slightly better description. The generating function for the number of individuals per sample when patchiness is included is then

$$G_{\chi}(s|k,p,\lambda) = \int_{o}^{\infty} d\mu\ f(\mu|k,p)e^{-\mu(1 - g_{\chi}(s|\lambda))} \quad . \tag{8}$$

Using the distribution function for the gamma distribution,

$$f(\mu|k,p) = \frac{\mu^{(k - 1)}p^{-k}}{\Gamma(k)}\ e^{-\mu/p} \tag{9}$$

($\Gamma(k)$ is the gamma function), we find

$$G_{\chi}(s|k,p,\lambda) = \left(1 + p - \frac{ps}{\lambda + s(1 - \lambda)}\right)^{-k} \quad . \tag{10}$$

The probabilities are (equation 4)

$$P(x) = \frac{1}{(1 + p)^{k}\lambda^{x}} \sum_{j=1}^{x} \binom{x - 1}{j - 1} \frac{(\lambda - 1)^{x-j}}{j!} \frac{\Gamma(j + k)}{\Gamma(k)} \frac{1}{(1 + 1/p)^{j}} \quad . \tag{11}$$

For ease of reference, we will call this distribution the regionalized Poly-Aeppli (RPA) distribution. Its mean and variance are

$$m = kp\lambda \quad ,$$

$$V = \frac{m^2}{k} + m(2\lambda-1) \qquad (12)$$

This distribution has three parameters:

k = patchiness parameter

p = scale parameter

λ = mean cluster size.

If λ is known then

$$k = \frac{m^2}{V - m - 2m(\lambda - 1)} \quad , \qquad (13)$$

which is similar to the negative binomial patchiness index except that correction for small scale clustering is made. In fact, for the limit $\lambda \to 1$ (no clustering), the RPA reduces to the negative binomial as may be seen from either the generating function or the probability distribution. For the case, $k \to \infty$ (no patchiness) the RPA becomes the Poly-Aeppli distribution and in both limits, it reduces to the Poisson. The RPA distribution is a very general distribution which we hope can accomodate many experimental situations.

To estimate parameters, it is best to design an experiment so that several samples are taken at a few stations and used to estimate λ and then several stations are sampled once to estimate k and p. Unfortunately, the data available to us is that normally collected by limnologists in large-scale lake experiments: it contained one sample from each station. Some direct evidence for cluster size was available, however, in most cases, it was necessary to estimate all three parameters

simultaneously. Because of this and because of the low number of stations sampled, we needed the most efficient estimates possible. We therefore used maximum likelihood to estimate all three parameters (Hoel Port, Stone, 1971). One likelihood equation reduces to $kp\lambda = \bar{x}$, the sample mean. The other two involve summations and are solved numerically by high speed computer (CDC 7600).

APPLICATION

We used our model to reduce data from synoptic studies made at Clear Lake, California from fall, 1969 through spring, 1972. Some of this data has been used to investigate the factors influencing nitrogen fixation in the lake (Horne, Dillard, Fujita and Goldman, 1972; Horne, Sandusky, and Carmiggelt, 1977), some to study algal ecology of the lake (Horne, Javornicky and Goldman, 1971; unpublished data of C.R. Goldman, P. Javornicky and A.J. Horne). These studies were conducted by one of two methods. In September, 1969 and in April, 1970 water samples were taken with a Van Dorn sampler (averaging the top 1/2 meter of the water column) at 32 locations in the 17,000 hectares of the lake surface. Algae were counted to genus in all cases, to species in most, at all of the stations. In spring, 1972, samples were taken from the same 32 stations using a tube sampler which averaged 5 meters of water column. This synoptic was repeated at approximately 3 day intervals for a total of six sampling periods ("supersynoptic") which included much of the spring bloom of Aphanizomenon. Algae were

counted to genus in most cases. In addition, water chemistry measurements (NO_3, D.O. etc.) were made as well as biological facotrs (nitrogen fixation, carbon 14, etc.). Details concerning station locations, data collection, and analytical procedures are given in the above mentioned articles.

Some amalgamation of the data was necessary prior to analysis. Rare species were excluded in cases where non-zero cell counts at ten or less stations were found or where the largest number of cells did not exceed five (we estimated patchiness for some species with smaller numbers). In the super-synoptic study, all 32 stations were sampled only on the 2nd and 4th sampling days. For the rest, data at only 15-16 stations was taken. Because this was too small a number to permit a reliable statistical analysis, data from the 1st and 3rd periods was dropped, data from the 5th and 6th periods was combined to make 32 samples with 16 from each period. In all, we had 43 sets of data representing 18 genera. These genera and mean numbers of counts are listed in Table 1.

To each set of data we fitted an RPA distribution, using the maximum likelihood method to compute values for the parameters. We tested goodness of fit using the likelihood ratio test statistic,

$$\chi^2 = -2 \sum_{i=1}^{m} O_i \log\left(\frac{E_i}{O_i}\right) , \qquad (14)$$

where E_i and O_i are the expected and observed numbers of events in the i^{th} cell. For large numbers of events, this statistic is distributed as χ^2 with m-r-1 degrees of freedom (r is the estimated number of parameters). For our data, we separated the counts into 5 "bins" (6-7 events in each

bin for those cases where the number of zero counts was not excessive). This separation was made by computer, after the parameters were estimated and without reference to the observations by making the bins of approximately equal size in probability space. Thus, we expect an equal number of observations in each bin (some fluctuations are unavoidable because the distribution is discrete). Three parameters are estimated; therefore, the goodness of fit statistic will be asymptotically distributed as χ^2 with one degree of freedom. We remark that 32 events is not asymptopia and fluctuations in our goodness of fit statistic caused by extreme good (ill) luck or by counting errors can be large. We compare results for our statistic with χ^2 tables, cautiously.

For comparative purposes, we also fit a negative binomial distribution to the data. This is the limiting case of large scale patchiness with no local clustering and we expect it to apply to certain groups of algae. Because the negative binomial can occur when there is no patchiness but local clustering, we checked two other distributions with two parameters, the discrete log-normal (patchiness without clustering) and the Neyman-A and Poly-Aeppeli distributions (clustering without patchiness). Results from this showed close correspondence with the discrete log-normal but not with the others, thus no ambiguity is present here. We used four bins for the negative binomial to make the goodness of fit statistic again, an asymoptic χ^2 with one degree of freedom. We compared the goodness of fit statistics for the two distributions to test for local clustering.

Comparison between the RPA distribution and the negative binomial also provides an interesting contrast in predicted values of the patchiness parameter. To study this parameter further, we fit a gamma distribution to the chemical-physical data. One of the parameters of the gamma distribution, the shape parameter, should be directly analogous to the algal patchiness parameter and allows some discussion of the influences of lake physics and chemistry on algal patchiness.

RESULTS

Local Clustering

When local clustering is present, we expect values of λ, the average number of units per cluster in the RPA distribution to exceed one. In the case of colonial algae where individual cell numbers are recorded, λ should be at least the average cells per colony. In Table 2, we list the cases where this occurred. Melosira is a chain diatom for which individual cells were counted. Aphanizomenon grows in bundles of large numbers of trichomes (as many as several hundred per bundle). When fixed with Lugols solution and shaken, these break up. The separation, however, is incomplete and several filaments often remain attached, thus, since individual filaments are counted, we expect clustering. The case for Anabaena and Oscillatoria is somewhat similar. While they do not occur in tightly attached bundles, many trichomes often stick together (sometimes to form large bundles) in numbers which depend upon lake conditions. Microcystis cells can

form colonies of a few to a few hundred thousand cells. Again, when fixed and shaken, it becomes difficult to decide what should be a separate entity since many of the clusters break apart. In fact, referring to Table 2, it is remarkable that while the Microcystis biomass changed considerably between the fall synoptic and the spring supersynoptic, the cluster density, $\bar{x}/\lambda$, was nearly the same (2.2 in the fall, 1.8 average in the spring). The change seemed to come almost entirely from a decrease in cluster size. This make biological sense for two reasons. First, fall aggregations occuring in a vigorous growth period are larger. Second, during such a period, the species is buoyant; thus separate aggregations may crowd each other, spatially.

As we remarked earlier, replicate sampling at a few stations is needed to provide reliable estimates of λ . Because that was not done in these experiments, values of λ found here are only approximate. Nevertheless, we see from Table 2 that all values appear to be substantially greater than one. Further if we compare χ^2 values between RPA and negative binomial (no clustering) in ten of twelve cases, RPA fit data more closely. This is further reflected in the totals. If we assume the totals are distributed as χ^2 with 13 degrees of freedom this difference is about 2.1 standard deviations ($P \leq .02$). That the total for RPA fails at 95% significance is not disturbing since nearly 8 points of 24 come from one set (Aphanizomenon, date 2). The total for the rest is well out of the critical region.

We expect little or no clustering for unicellular flagella and, referring to Table 3, we find strong evidence for this. Values for both λ and χ^2 differ markedly from those in Table 2. The negative binomial distribution fits better in most cases (with $\lambda = 1$, the RPA distribution and the negative binomial are the same – the extra parameter used by the RPA results in a larger χ^2 since one more bin is used). The total χ^2 values differ by 4.9 standard deviations ($P \sim o$).

Non-random spatial associations might occur for non-motile species at small distances if reproduction by fission is faster than diffusion or if a species grows epiphytically on other algae. Blue-green algae may be clustered because they can regulate their depth by producing gas vacuoles as seems to be the case for Microcystis. In Table 4 we list non-flagellated species where the proper unit (be it cell or filament) was counted. Referring to estimated values of λ, little clustering seems evident for the centric diatoms. For the green algae, clustering may occur. If association distances involved are between 0.5 m and 5 m, a difference should be noticeable between the two sampling methods. No clear difference is present and stronger conclusions concerning local associations await further experimentation.

Patchiness

Values for algal patchiness depend on the method used to estimate it. Our results for clustering indicate that it is best to use RPA distribution parameter for blue-green, non-motile green and all colonial

algae (where cells or subunits are counted). Since no evidence of local clustering was found for unicellular flagellates or for centric diatoms (when filaments were counted), we use the negative binomial distribution (i.e. the no-clustering limit of the RPA) to estimate patchiness for these algae. As large values of k carry large errors and as it is desirable to have an index which increases with increasing patchiness, we transform the RPA (negative binomial) parameter to

$$L = 1 + 1/k \qquad (15)$$

If no clustering is present and if the method of moments is used to estimate k, this parameter is the same as Lloyd's index (Lloyd, 1967). A value of $L = 1$ denotes a random distribution. Large values of L signify patchy distributions. In Table 1, we list values for patchiness. Although the maximum likelihood method does allow calculation of error in parameter estimation, we do not list the error estimates since for our sample numbers they are unreliable. Errors did tend to increase with increasing patchiness. An alternate estimate of the reliability of our index is obtained by comparing the 3 different periods of the supersynoptic. Only some between-day variability is present, notably in those species with low mean counts. Differentiation between species is clearly feasible and we can use it to understand some basic mechanisms causing patchiness.

Experimental methodology is an important consideration in measuring patchiness. If a lake is stratified and only part of the photic zone is sampled, increased variability (hence an erroneously high estimate of patchiness) may be observed for non-motile organisms whose vertical

position is dependent on sieche waves (Platt and Conover, 1971). Clear Lake is well mixed during spring and fall blooms when the sampling was done and the only group whose patchiness might depend on sampling strategy is the blue-green algae. In Table 1, we see both Aphanizomenon and Microcystis show considerably greater patchiness when sampled using a Van Dorn sampler (0-.5 m) than when sampled with a depth integrating tube. For reasons discussed in the next paragraph, the result for Microcystis may be coincidental. Aphanizomenon, however, seems to show substantial difference.

A second consideration in sampling strategy is station bias since all station positions may not be equivalent. In Fig. 1a, we see that Microcystis is growing only in the Lower and Oaks Arms due most likely to ammonia releases related to sediment content. While the bloom eventually spread to the Upper Arm, the degree of patchiness is inflated because of station differences. Another example of deterministic influences is Peridinium (Fig. 1b) which grows best in the Upper Arm due to favorable inflow and temperature conditions (Horne, Javornicky and Goldman, 1971). Other species showed little or no evident bias.

Spatial patterns of nutrients, chemical and physical factors are closely linked with phytoplankton patterns. In the supersynoptic study, many of these quantities were measured. To compare their patchiness with that of algae, we fit a gamma distribution to those data sets where enough measurements were made to yield reliable results. In Table 5, we list values for the shape parameter, the continuous

analogue of the patchiness parameter for phytoplankton. For easy comparison, we again transform to $L = 1 + 1/k$. Major nutrients, NH_4, PO_4, and NO_3 exhibited patchiness comparable to the algae. Nitrate is the most patchy of the three; it is interesting conjecture that many of the nitrate limited species are more patchy than those limited by other factors. Biological factors, C_{14}, acetylene reduction (nitrogen fixation) and Chl. a are approximately the same as Aphizomenon, the dominant algae, as would be expected. Physical parameters are less patchy. Temperature fluctuations are underestimated since the zero temperature is arbitrary (hence, by a change of zero, we could change the variance to mean ratio) and since phtoplankton-temperature relationships are complicated.

Patchiness of a particular species changes from season to season. For a given species of algae we conjecture that environmental fluctuations occuring during a dominant period are less likely to affect growth than those occuring in a subdominant one. Thus, if we compare patchiness of a particular species at two different seasons, we expect this species to be less patchy during the period for which its biomass is greater. Using Table 1, we see that Aphanizomenon, Rhodomonas, and Cryptomonas Marssonii increased in numbers between fall and spring and showed decreased patchiness. Scenedesmus was unchanged in both numbers and patchiness. Cryptomonas Reflexa showed increased patchiness but no change in numbers, although in both seasons it was rare and estimates of its patchiness could be unreliable. Comparisons between synoptic

and supersynoptic are supportive for Cyclotella and Cryptomonas when combined species are considered (comparisons for blue-greens are not meaningful because of sampling differences).

A major ecological factor influencing patchiness, here, is what can be called species sensitivity or niche width. A partially mixed lake presents a variety of habitats of different chemical-physical makeup. Algae which are extremely adaptable should do well in many of them and thus be found in similar quantity at all stations (random distribution). Algae which are extremely sensitive to environmental fluctuations will manifest greater patchiness since their expected numbers depend greatly on local water conditions. We can see some evidence for this conjecture from Table 1. Blue-green algae which are adaptable to a variety of conditions are less patchy then either flagella or centric diatoms (not enough species of green algae are available for a comparison, although the ones shown show little patchiness).

An interesting way to further explore the relationship of nitch width and patchiness comes by studying combinations of algal species. If two closely related species (say, two cryptomonads of about the same size) are summed, we expect decreased patchiness since the niche occupied by the combination is larger than for either separately. Theoretically, this concept is built into the RPA distribution since if two independent variables which have a gamma distribution with the same scale parameter are combined, their sum will be gamma distributed with the shape

parameter equal to the sum of the individual shape parameters. If the two variables are linearly dependent, their sum will be gamma distributed with the same shape parameter. For partially dependent variables, the result lies (approximately) between. This carries over (approximately) to discrete distributions which are based on a gamma distribution as are the RPA and negative binomial. In short, if the niche is not expanded by combining two species the patchiness parameter will remain unchanged (for our case, it can be estimated by the weighted average of the two individual parameters if they differ). If it is increased due to partial or total variable independence, the patchiness decreases. In Table 6, we list several combinations of related species which had about the same scale parameter. As expected, combinations of species have combined patchiness close to the dependent case (L_D). Values of patchiness obtained for combinations of two phyla are closer to the independent case (L_I).

DISCUSSION

Conceptually, experimental studies of patchiness can be categorized as deterministic or stochastic. A deterministic approach has as its objective, measurements of size, shape, density and/or evolution of plankton patches. It is useful for qualitative or semiquantitative descriptions of mesoscale patches (Bainbridge, 1957) caused by regional environmental conditions favorable for production (Cushing, 1955) or

favoring particular species (Holland and Beeton, 1972). These large scale patches are thought to be stable (Kierstead and Slobodkin, 1953). A deterministic approach is also useful for descriptions of vertically stratified plankton.

Stochastic descriptions apply in two cases: (1) where insufficient data is present to permit characterization of a patch or (2), where no fixed differences in environmental factors occur among sampling points. For these studies, spatial variability is often of primary interest. Sizes are "measured" not of individual patches but as correlations observed as a function of distance (Platt and Denman, 1975; Powell et al., 1975) or as fitted parameters of a hypothesized frequency distribution (Fasham, Angel and Roe, 1974).

Our description of patchiness was stochastic. Because our samples were not made at stations with clear spatial relationships, we did not attempt to include patch size as a variable. To describe transect data, our model could be expanded. However, inclusion of fallacious assumptions concerning patch size or shape could lead to an untrustworthy analysis. A better approach is to use our distributions as a "normalizing" transformation to use prior to other analyses. Power spectral analysis or other techniques based on normal distributions could then be used with greater accuracy. For the RPA distribution, this transformation is

$$f(x) = \frac{2}{\sqrt{L-1}} \sinh^{-1}\sqrt{\frac{x(L-1)}{(2\lambda-1)}} \quad , \qquad (16)$$

where $L = 1 + 1/k$, as before. For $\lambda \to 1$, this becomes identical with the negative binomial transformation. For large sample numbers, it reduces to log(x).

Local Associations

A major problem with many quantitative studies on patchiness has been the one of deciding upon a fundamental unit (Pielou, 1969). In terrestrial and in aquatic benthic studies, plants aggregate (cluster) because of limited dispersal ability. Many phytoplankton species also occur in colonies or filaments (aggregates). In principle, this problem is avoided by counting aggregates. In practice, it is often difficult or impossible to decide which cells belong to which aggregate. This is especially true for loosly attached species, for example, most bloom forming cyanophytes, which break up when preserved and shaken. Further, if plankton are vertically stratified, aggregates may be concentrated, physically, and the unit for counting will be some ill-determined "super-aggregate". We provided for local associations in our model. Because the data which we analyzed was recorded regionally, our conclusions about local associations are indirect. With the exception of buoyancy controlled cyanophytes, we found little evidence for spatial associations, locally. This is in contrast with work by several authors (Cassie, 1963, and references therein). The discrepancy is easily understood since Clear Lake is shallow and well mixed (vertically) at the time of sampling. Most of the non-random spatial associations observed in the literature

occur for depth stratification of stronger swimming marine zooplankton or dinoflagellates or for passive organisms in stratified systems (Cassie, 1960). In some cases, local association was significant but not pronounced. Barnes and Marshall (1951) used contagious distributions to fit samples taken from a drifting boat. Their frequency spectra are all unimodal and it appears that our distribution describing local associations (Poly-Aeppli) would fit them. Using their means and variances, we estimated values of λ for the various species and hauls. Our average value for all their listed species was $\lambda = 1.44$. This is greater than the values we found for most unicellular algae. The difference is probably due to sampling differences and to the greater salinity (pressure) gradients observed by these authors.

In addition to the intrinsic interest in local associations, proper account of them is necessary for a good estimate of regional associations. If we underestimate clustering, we will overestimate patchiness. A quantitative description of this may be found using the RPA distribution. The classic patchiness index (Fisher, 1950; Cassie, 1963) is the difference between variance and mean. From Equations (13) and (15).

$$\frac{V-m}{m} = m(L - 1) + 2(\lambda - 1) \quad . \tag{17}$$

Both regional overdispersion $(L > 1)$ and local overdispersion $(\lambda > 1)$ contribute to a large variance. This explicit partitioning of variance into local and regional spatial patterns is an interesting conceptual feature of the RPA distribution.

Regional Associations

When solitary plankters are counted and when local mixing diminishes local associations, no clustering is expected, thus the RPA distribution reduces to the negative binomial ($\lambda = 1$). Other applications of the negative binomial to regionally sampled plankton populations have been successful. Colebrook (1960) approximated a hydrodynamical model of lake mixing with a negative binomial and found good fits to observed numbers of zooplankton. Cassie (1962) found comparable fits for a negative binomial and truncated log-normal. He proposed use of the discrete log-normal distribution, however, he was unable to use it effectively since the evaluation of probabilities necessitated a numerical approximation to an integral. We were able to accurately evaluate this integral using Gaussian quadratures and a high speed computer. Comparison of fits to our observed frequencies of unicellular flagella for the discrete log-normal and the negative binomial distributions revealed close similarity between the two with slightly smaller χ^2 values for the latter. Both distributions are derived by assuming that patchiness (overdispersion) is due to regional fluctuations in water mass chemistry and expected plankton numbers not by assuming random spatial distributions of clusters as we did for local fluctuations.

Patchiness indices in general and our index in particular are derived with reference to some standard of randomness. Because of uncertainties in values of this index caused by sampling procedures and hydrodynamic mixing, perhaps a better concept is "relative patchiness" where one looks simultaneously at species patchiness relative to other species or

nutrients or at patchiness of the same species sampled on different days but measured using the same technique. Such comparisons were useful for us. We found similar patchiness of limiting nutrients and phytoplankton. Our results linked relative phytoplankton patchiness to relative dominance (same species, different seasons) and to species sensitivity or niche width (different species, same sampling period). Our results also showed that relative patchiness decreased when similar species were combined (enlarged niche). This has possible applications to survey experiments where the effects of an environmental disturbance are measured for a group containing several species (Niebold, 1977) or where spatial fluctuations in a biological indicator such as chlorophyll are measured (Powell et al., 1975; Platt and Denman, 1975). For the latter case, our results suggest the degree of plankton spatial organization may be underestimated when several species are present. In future work, we hope to investigate more systematically these relationships between patchiness and algal ecology.

REFERENCES

Bainbridge, R. 1957. The size, shape, and density of marine phytoplankton concentrations. Biol. Rev. 32:91-115.

Barnes, H. and S.M. Marshall 1951. On the variability of replicate plankton samples and some applications of contagious series to the statistical distribution of catches over restricted periods. J. Mar. Biol. Assoc., U.K. 30:233-263.

Bernhard, M. and L. Rampi 1965. Horizontal microdistribution in the Ligurian Sea. Proc. Mar. Biol. Symp. (5th) Goteburg, pp.13-24.

Bittel, L., S. Grimm, and S. Ochocki. 1965. The horizontal distribution of zoo-plankton in the southern part of Lake Jeziorak. Zesz. Nauk. UMK. Biol. 13:21-46.

Cassie, R.M. 1959. An experimental study of factors inducing aggregation in marine plankton. N. Z. Jl. Sci. 2:339-365.

Cassie, R.S. 1960. Factors influencing the distribution pattern of plankton in the mixing zone between oceanic and harbour waters. N. Z. J. Sci. 3:26-50.

Cassie, R.M. 1962. Frequency distribution models in the ecology of plankton and other organisms. J. Anim. Ecol. 31:65-92.

Cassie, R.M. 1963. Microdistribution of plankton. in Barnes, H. and G. Allan (Eds.), Oceanography and Marine Biology 1:223-252.

Colebrook, J.S. 1960. Plankton and water movements in Windemere. J. Anim. Ecol. 29:217-240.

Cushing, D.H. 1955. Production and a pelagic fishery. Fish. Invest. Lond. Ser. II, 18 No. 7:104 pp.

Fasham, M.J.R., M.V. Angel and H.S.J. Roe 1974. An investigation of the spatial pattern of zooplankton using the Longhurst-Hardy plankton recorder. J. Exp. Mar. Biol. Ecol. 16:93-112.

Feller, W. 1950. An introduction to probability and statistics. Wiley & Sons.

Fisher, R.A. 1950. The significance of deviations from expectation in a Poisson series. Biometrics. 1:17-24.

Hill, M.O. 1973. The intensity of spatial pattern in plant communities. J. Ecol. 61:225-235.

Hoel, P., S. Port and C. Stone. 1971. Introduction to Statistical Theory. Houghton Mifflin Co., Boston.

Holland, R.E. and A.M. Beeton 1972. Significance to eutrophication of spatial differences in nutrients and diatoms in Lake Michigan. Limnol. and Oceanogr. 17:88-96.

Horne, A.J., P. Javornicky and C.R. Goldman 1971. A freshwater red tide in Clear Lake, California. Limnol. Oceanogr. 16:684-689.

Horne, A.J., J.E. Dillard, D.K. Fujita and C.R. Goldman 1972. Nitrogen fixation in Clear Lake, California. II. Synoptic studies on the autumn Anabaena bloom. Limnol. Oceanogr. 17:693-703.

Horne, A.J. and R.C. Wrigley 1975. The use of remote sensing to detect how wind influences planktonic blue-green algal distribution. Vert. Int. Vert. Limnol. 19:784-791.

Horne, A.J., J.C. Sandusky, and C.J.W. Carmiggelt 1977. Nitrogen fixation in Clear Lake, California III. Supersynoptic studies on the spring Aphanizomenon bloom. Limnol. Oceanogr. in press.

Kierstead, H. and L.B. Slobodkin 1953. The size of water masses containing plankton blooms. J. Mar. Res. 12:141-147.

Lloyd, M. 1967. Mean crowding. J. Anim. Ecol. 36:1-30.

Lund, J.W.G. 1965. The ecology of the freshwater phytoplankton. Biol. Rev. 40:231-293.

Margalef, R. and M. Estrada 1971. Simple Approaches to a pattern analysis of phytoplankton. Invest. Pesq. 35:269-297.

McAlice, B.J. 1970. Observations on the small-scale distribution of estuarine phtoplankton. Mar. Biol. 7:100-111.

Mullin, M.M. and E.R. Brooks 1976. Some consequences of distributional heterogeneity of phytoplankton and zooplankton. Limnol. Oceanogr. 21:784-796.

Neyman, J. 1939. On a new class of contagious distributions applicable in entomology and bacteriology. Ann. Math. Stat. 10:35-37.

Newbold, J.D. 1977. The use of benthic macroinvertebrates as indicators of logging impact on streams with an evaluation of buffer strip effectiveness. Ph.D. thesis, University of California, Berkeley.

Patil, G.P. 1968. Random counts in scientific work. Vol. 2: Random counts in biomedical and social sciences. Penn. State Univ. Press.

Peilou, E.C. 1969. An introduction to mathematical ecology. Wiley Interscience, New York.

Platt, T. and R.J Conover. 1971. Variability and its effect on the 24 hour chlorophyll budget of a small marine basin. Marine Biology 10:52-65.

Platt, T. 1972. Local phytoplankton abundnace and turbulence. Deep-Sea Res. 19:183-187.

Platt, T. and C. Filion 1973. Spatial variability of the productivity: biomass ratio for phytoplankton in a small marine basin. Limnol. and Oceanogr. 18:743-749.

Platt, T. 1974. The physical environment and spatial structure of phytoplankton populations. Proc. 6th Liege Coll. on Ocean Hydrodynamics, Mem. Soc. Roy. Sci. Liege. 9-19.

Platt, T. and K.L. Denman 1975. Spectral analysis in ecology. prepared for Annual Review of Ecology and Systematics, Vol. 6.

Polya, G. 1931. Sur quelques points de la theorie des probabilities. Anales de l'Institut Henri Poincare 1:117-161.

Powell, T.M., P.J. Richerson, R.M. Dillon, B.A. Agee, B.J. Dozier, D.A. Godden and L.O. Myrup 1975. Spatial scales of current speed and phytoplankton biomass fluctuations in Lake Tahoe. Science. 189: 1088-1090.

Richerson, P., R. Armstrong and C.R. Goldman. 1970. Contemporaneous disequilibrium, a new hypothesis to explain the 'paradox of the plankton'. Proc. Nat. Acad. Sci. 67:1710-1714.

Riley, G.A. 1976. A model of phytoplankton patchiness. Limnol. Oceanogr. 21:873-880.

Stavn, R.H. 1971. The horizontal-vertical distribution hypothesis: Langmuir circulations and Daphnia distributions. Limnol. Oceanogr. 16:453-466.

Steele, J.H. 1974. Spatial heterogeneity and population stability. Nature. 248:83.

Wrigley, R.C. and A.J. Horne 1975. Surface algal circulation patterns in Clear Lake by remote sensing. NASA TMS-62, 451.

TABLE AND FIGURE CAPTIONS

Table 1: Mean frequencies and patchiness indices for phytoplankton. Patchiness values (Equation 15) are italicized. The dates corresponding to the letters are September, 1969 (F); April, 1970 (S); 2 June 1972 (2) and 12 June, 1974 (4). The data for date (5) was compiled from samples taken on 15 June and 19 June.

Table 2: Cluster sizes and χ^2 values for colonial or filamentous phytoplankton species for which the counting unit was a cell or subaggregate of the cluster of cells. Dates are the same as for Table 1. To avoid excessive computer costs, values of λ for A. flos-aquae denoted by an asterisk were set equal to 7.0, the value obtained for the September, 1969 data. The average value for λ is a geometrical average.

Table 3: Cluster sizes and χ^2 values for unicellular flagella. Dates are the same as for Table 1.

Table 4: Cluster sizes and χ^2 values for non-motile species where the proper unit (cell or filament) was counted. Dates are the same as for Table 1.

Table 5: Patchiness of major nutrients, physical and biological factors. Dates are as in Table 1 except for C_{14} (asterisk) which is 6 June and 12 June, 1972 combined.

Table 6: Patchiness of phytoplankton groups. L_{A+B} is the patchiness parameter (Equation 15) obtained by fitting a negative binomial distribution to the total of observed frequencies for species A and B. L_I is the expected value if the two species were independent, $L_I = 1 + 1/(k_A + k_B)$. L_D is the expected value if species were dependent, $L_D = 1 + (P_A + P_B)/(P_A k_A + P_B k_B)$.

Figure: Spatial distribution of Microcystis aeruginosa and Peridinium penardii in Clear Lake. Data taken in September, 1969 (Microcystis) and April, 1970 (Peridinium).

TABLE 1

PHYLUM	GENUS	DATES F		S		2		4		5	
CYANOPHYTA	Anabaena circinalis	560.	*2.1*								
	Aphanizomenon flos-aquae	22.	*5.8*	680.	*3.0*	700.	*2.2*	380.	*1.6*	760.	*1.6*
	Microcystis spp.					11.	*1.8*			7.9	*3.1*
	M. aeruginosa	1100.	*4.2*								
	Oscillatoria spp.					6.4	*1.9*	5.2	*1.1*	2.8	*1.0*
CHRYSOPHYTA	Coscinodsicus spp.					.45	*5.0*			.84	*4.4*
	Cyclotella spp.					1.0	*4.3*			1.2	*9.7*
	C. sp.			17.	*4.4*						
	C. atomus			530.	*1.8*						
	Mallomonas spp.			3.5	*1.4*						
	Melosira granulata v. granulata	87.	*1.5*								
	M. italica			.36	*1.3*						
DINOPHYCAE	Peridinium penardii			4.8	*17.7*						
CRYPTOPHYCEAE	Chroomonas sp.			290.	*1.9*						
	Cryptomonas spp.	2.3	*1.2*			3.1	*5.5*	.64	*5.2*	.41	*4.4*
	C. erosa			5.8	*2.6*						
	C. marssonii			3.6	*2.8*						
	C. reflexa			1.9	*6.0*						
	Flagellates					2.6	*2.4*	7.8	*1.7*	.84	*8.1*
	Rhodomonas pusilla	23.	*1.9*	100.	*1.5*						
CHLOROPHYTA	Chlamydomonas spp. and zoospore			75.	*2.5*						
	Monoraphidium contortum			4.7	*2.1*						
	Oocystis lacustris	.47	*1.0*								
	Scenedesmus spp.	2.1	*1.1*	1.6	*1.1*						
	Schroederia spp.					4.4	*1.1*	3.5	*1.1*	1.4	*1.0*

TABLE 2

Species	Date	χ^2 RPA	χ^2 NB	Best Fit	λ
Anabaena circinalis	F	0.53	0.73	RPA	20.
Aphanizomenon flos-aquae	F	0.33	2.92	RPA	7.0*
Aphanizomenon flos-aquae	S	0.49	1.27	RPA	7.0
Aphanizomenon flos-aquae	2	7.66	7.66	-	7.0*
Aphanizomenon flos-aquae	4	3.01	1.50	NB	7.0*
Aphanizomenon flos-aquae	5	1.02	1.01	NB	7.0*
Melosira italica	S	0.13	0.88	RPA	3.4
Microcystis aeruginosa	F	4.23	5.43	RPA	500.
Microcystis spp.	2	0.01	3.71	RPA	11.1
Microcystis spp.	5	0.76	0.84	RPA	3.0
Oscillatoria sp.	2	0.70	1.60	RPA	2.4
Oscillatoria sp.	4	0.80	1.60	RPA	6.1
Oscillatoria sp.	5	4.46	5.47	RPA	2.2
Totals		24.13	34.62	RPA	8.1

TABLE 3

Species	Date	χ^2 RPA	χ^2 NB	Best Fit	λ
Chroomonas sp.	S	1.56	1.29	NB	3.0
Chlamydomonas spp.	S	6.73	0.72	NB	1.0
Cryptomonas spp.	2	7.93	0.33	NB	1.0
C. erosa	S	7.59	2.33	NB	1.0
C. marssonii	S	5.53	7.21	RPA	1.0
C. reflexa	S	2.00	0.07	NB	1.3
Flagellates	2	2.01	0.08	NB	1.04
Flagellates	4	2.93	4.29	RPA	1.0
Mallomonas spp.	S	3.76	2.65	NB	1.61
Rhodomonas pusilla	F	5.47	2.89	NB	1.0
Rhodomonas pusilla	S	1.08	1.94	RPA	1.0
Total		46.59	23.80	NB	1.2

TABLE 4

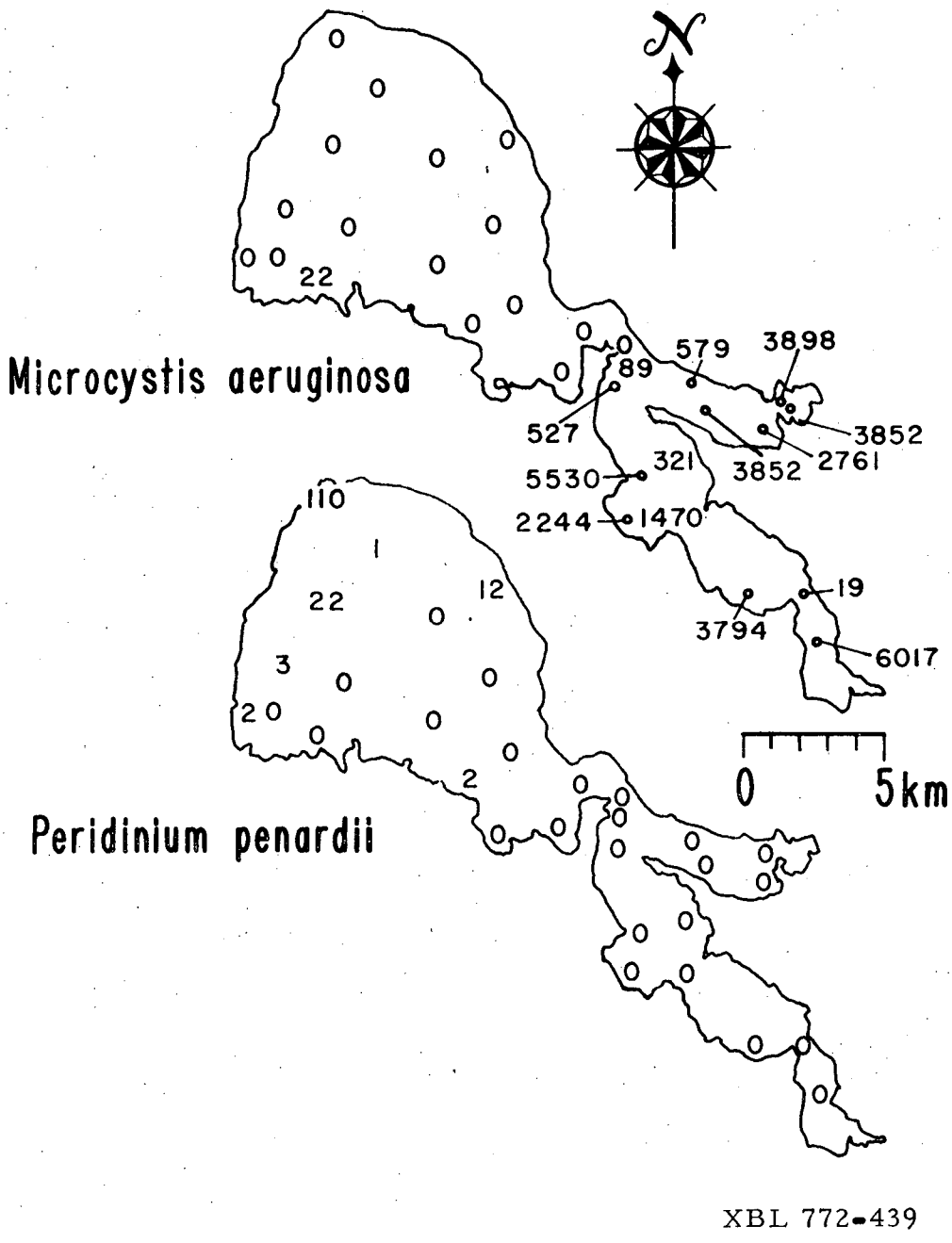
Species	Date	χ^2 RPA	χ^2 NB	Best Fit	λ
Cyclotella spp.	2	1.58	0.18	NB	1.1
C. sp.	S	2.36	0.79	NB	1.0
C. atomus	S	4.14	1.89	NB	1.8
Melosira granulata	F	1.06	3.77	RPA	1.0
Monoraphidium contortum	S	1.77	1.65	NB	1.5
Schroederia spp.	2	6.06	6.07	RPA	1.9
Schroederia spp.	4	1.84	5.84	RPA	2.6
total		18.81	20.19	RPA	1.5

TABLE 5

Date	2	4	5 + 6
Temperature	1.0	1.0	1.0
Turbidity	1.2	1.1	
NH_4		2.1	1.8
NO_3		2.1	2.9
PO_4		1.5	1.3
acetylene reduction	1.8	1.9	1.6
C_{14}		1.7*	1.4
Chl. a	1.8	1.7	1.7

TABLE 6

Combination Type	A	B	Date	Cluster Frequency		Patchiness		
				P_A	P_B	L_I	L_{A+B}	L_D
species	Cryptomonas erosa	Cryptomonas reflexa	S	9.3	9.5	2.2	2.6	3.4
species	Cryptomonas erosa	Cryptomonas marssoni	S	9.3	6.6	1.8	2.4	2.7
genus	Rhodomonas pusilla	Chroomonas sp.	S	51.0	255.	1.3	1.7	1.8
genus	Cyclotella spp.	Coscinodiscus spp.	2	3.4	1.8	2.8	3.3	4.5
genus	Cyclotella spp.	Coscinodiscus spp.	5	10.9	2.9	3.5	3.8	7.6
family	Oocystis lacustris	Scenedesmus spp.	F	2.1	17.4	3.9	6.6	8.7
phylum	Flagellates	Cryptomonas spp.	2	3.9	14.0	2.1	2.6	4.1
phylum	Flagellates	Cryptomonas spp.	4	5.3	2.7	1.6	1.7	1.9
phylum	Flagellates	Cryptomonas spp.	5	5.9	1.4	3.3	3.3	6.8



XBL 772-439

Fig. 1

This report was done with support from the United States Energy Research and Development Administration. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the United States Energy Research and Development Administration.

TECHNICAL INFORMATION DIVISION
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720