PROCEEDINGS OF THE CALIFORNIA ACADEMY OF SCIENCES

Volume 56, No. 13, pp. 146–155, 9 figs. April 29, 2005

Marrying Taxonomy and Ecology: An Attempt

Eugene F. Stoermer

Michigan Herbarium and School of Natural Resources and Environment,
University of Michigan, Ann Arbor, MI 48109-1115;*Email*: *stoermer@umich.edu*

Freshwater diatoms present an interesting challenge in an era when biodiversity is becoming a major concern. Although tremendously useful indicators of ecological conditions, past and present, lack of taxonomic knowledge limits the potential of ecological interpretation. At the same time the ecological studies that are carried out provide less than optimal feedback to the taxonomic literature. I suggest that appropriate use of available computer-based technologies can integrate these fields to the benefit of both. I further outline the approaches taken in an early and primitive attempt to accomplish this goal, the benefits derived, and the mistakes made and inadequacies of our effort at that time. Thoughtful application of technologies now available has the potential to further integrate studies and expand eventual understanding.

The following is a discussion of an attempt to marry the fundamental approaches of systematic practice to ecological studies. The tools and approaches used are, in retrospect, quite primitive, but there is an underlying logical framework that applies to all such endeavors, in taking on any problem at any time. I hope that discussing the way we attempted to solve problems common to any taxonomically based diatom study, what worked and what didn't, and the mistakes made, will be of some value to current investigators.

I should hasten to explain that the "we" in the previous paragraph is used advisedly. I am not a programmer, so much of the development and implementation was done by other people, better equipped to deal with the intricacies of programming than I. I thank the late Dr. Vincent Noble and Dr. Edward Johnston (Johnston and Stoermer 1976) for enlightening discussions of logical structures appropriate for human — computer interactions. The initial programming was done by Dr. J.K.C. Huang and the system was brought to its most advanced state largely through the efforts of Theodore and Barbara Ladewski (Ladewski and Stoermer 1973; Sicko-Goad et al. 1977). Numerous helpful comments and suggestions were also made by many technical staff and students, which materially helped shape the project.

The problem

In the mid 1960s I was a young investigator faced with the rather intimidating problem of investigating the algal flora of the Laurentian Great Lakes. At the time, severe eutrophication problems were apparent in many regions of these lakes (Beeton 1965, 1969). Because of the Great Lakes' tremendous value to the economies of the United States and Canada, considerable resources were available for studies related to water quality. Many of the practical problems that beset the lakes at that time were directly related to algae. Taste and odor problems caused by diatoms in the spring (Vaughn 1961, 1962) and cyanophytes in the summer and fall (Stoermer and Stevenson

1980; Bierman and Dolan 1981; Stoermer and Theriot 1985). *Cladophora* was a nuisance in many regions of the lakes (Wolfe and Sweeney 1980) and generally unpleasant obnoxious conditions were present in many areas. Lake Erie, in particular, became a cause célèbe of the environmental activism of the day, and was widely reported in the common press to be a "dead lake." This was somewhat problematic to biologists, as the actual problem was over-production, which eventually led to de-oxygenation of the bottom waters in certain areas of the lakes, creating so called "dead zones" where benthic invertebrates were periodically exterminated. In retrospect, the problems of the 1960s and 1970s were only the most recent in a long history of environmental catastrophes, such as epidemics of water-borne diseases (Beatty 1982; Bonner 1991) that devastated communities that drew drinking water from the lakes. For example, the great cholera epidemic of 1854 was estimated to have killed five percent of the total population of the city of Chicago. Collapse of native fish stocks began soon after western settlement of the region (Smith 1972), and culminated in total extermination of some native stocks by 1950 (Beeton 1969) and introduction of many exotic fish species.

One would rationally suppose such a valuable, but clearly damaged, ecosystem would have received careful and comprehensive study, especially considering the large number of well-known academic institutions in the region. Unfortunately, this was not the case. The ecological history of the Great Lakes, in many respects, provides a sterling example of precisely the wrong way to approach management of a large and complex ecosystem. Each successive crisis generated a wave of "directed research" centered on the apparent problem and to a lesser extent, if at all, on its root causes. "Charismatic vertebrates," in this case fish, were the initial center of attention, and lesser attention and resources were devoted to the rest of the biota or to chemical and physical factors of the environment.

In the case of diatoms, early (in the North American context) exploratory studies were carried out by J.W. Bailey in 1839, first mentioned in 1842 (Bailey 1842a, 1842b), and sent to C.G. Ehrenberg, who more formally published them in his monumental works (Ehrenberg 1845, 1854). These collections are still maintained at the Museum für Naturkunde, Humboldt-Universität zu Berlin, and have been used in more recent studies of the Great Lakes diatom flora (Stoermer and Ladewski 1982). Early pollution studies, particularly in the area of Chicago (Thomas and Chase 1887) and Cleveland (Vorce 1881, 1882) produced collections which are still available, but the majority of taxonomic work undertaken was either un-vouchered, or the material resulting from the study has been lost. For example, studies on early fisheries declines included some work on diatoms (e.g., Ward 1896; Thompson 1896) but we have never been able to locate any of these collections.

Thus, from the beginning it was apparent that the type of supporting references and materials generally assumed to be available to ecological studies were lacking. Although this problem is obvious in the Great Lakes case, it applies to the majority of studies attempting to use diatoms as ecological indicators, as I have argued elsewhere (Stoermer 2001).

APPROACH

Collections

Early on I determined that it was absolutely necessary to maintain a consistent and reasonably well ordered reference collection. It was clear that the available literature of the time was grossly insufficient to support repeatable identifications, so the availability of a reference standard was essential. Maintenance of vouchers, once a routine part of good scientific practice, has largely been abandoned in ecological studies. Logically, it is still necessary for studies involving lesser-known organism groups, and certainly should be a requirement for studies involving diatoms. It is sometimes argued that maintaining collections is "too expensive" for the competitive world of ecological funding. In a reasonable and logical world the functions of developing a comprehensive taxonomy might be separated, as they are in most large organisms, but this was not the case at the time I began. Although it has become much easier in recent years, due to general recognition of the biodiversity crisis, in the 1960s and 1970s it was virtually impossible to obtain direct funding for taxonomic studies of microscopic eukaryotes.

In our case, I simply made the decision that studies from our lab would be supported by vouchers, as a minimum standard of scientific practice. Our collections are in the form of lots, numbered consecutively. Each lot consists of raw material, cleaned material, and one or more slides. In some cases, we have accepted slides from other investigators and integrated them into the collection without other material, but this is a compromise to be avoided if at all possible. Because we operated primarily from ships, locality information consists of latitude and longitude and brief habitat and collection method descriptors. With the current availability of global positioning system (GPS) apparatus, there is now no excuse not to substitute this unambiguous information for references to inconstant physical landmarks and place names. In the better systems available, it is also possible to directly transcribe information electronically, avoiding the inevitable mistakes introduced by hand transcription.

Index and Pictorial Reference

When working on a system such as the Great Lakes it is easy to escape the illusion that appropriate names for all diatoms encountered exist in the literature, or the equally pernicious assumption that all names in the literature reflect biological reality. For that reason, we have always treated diatom names as entirely arbitrary. Thus, a nomenclaturally correct binomial is quite acceptable but, in our system, an arbitrary name (e.g., aff. *Navicula ambigua*) or a numerical designation (e.g., *Nitzschia* 343) is equally acceptable, *if it is supported by an adequate illustration and voucher specimen*. This, of course, is a compromise, recognizing the fact that it is not possible to resolve all taxonomic questions while conducting ecological studies, which furnished support for our lab at the time the system was instituted. In order to keep internal consistency, but avoid the extra time and effort necessary to directly compare specimens under a microscope, we resorted to a photographic archive. An illustration of the file used is shown in Figure 1. The elements are an epithet (upper left), one or more photographs (upper right), the dimensions of the specimen(s) (center) and coordinates of their location on a slide (in parentheses) taken from a particular microscope indicated by the letter following. Pictorial representations of specimens circled on a slide, and location of specimen(s) within a particular circle (lower left), photo magnification (lower center) and the collection number (lower right) are also provided. In our original system, additional notes were written on the back of

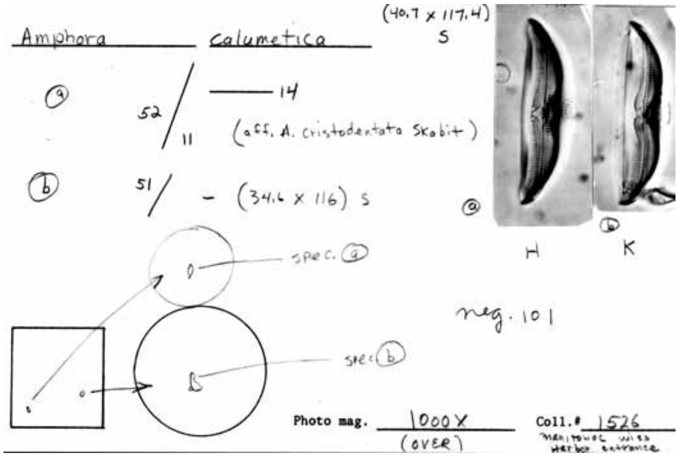
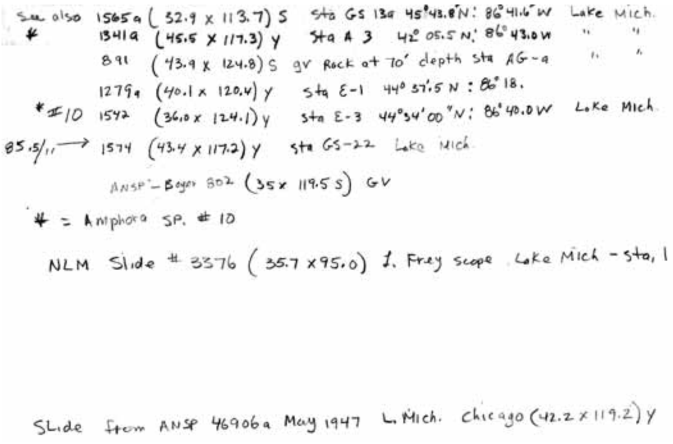


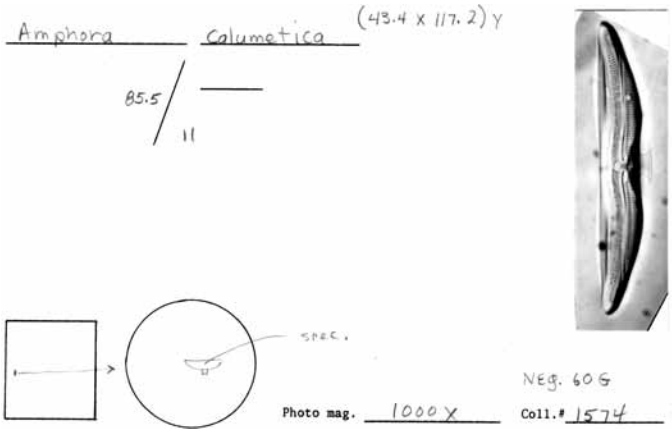
FIGURE 1. Example of card image used for specimen location and identification. See text for explanation.



See also 1565a (52.9 x 113.7) S sta GS 13a 45°43.8'N: 86°41.6' W Lake Mich.
\# 1341a (45.5 x 117.3) y Sta A 3 42° 05.5 N: 86° 43.0 W " "
891 (43.9 x 124.8) S gv Rock at 70' depth sta AG-a " "
1279a (40.1 x 120.4) y sta E-1 44° 57.5 N : 86° 18.
*#10 1542 (36.0 x 124.1) y sta E-3 44°54'00"N: 86° 40.0 W Lake Mich
85.5/11 → 1574 (43.4 x 117.2) y sta GS-22 Lake Mich
ANSP – Boyer 802 (35 x 119.5 S) GV
\# = Amphora sp. # 10
NLM Slide # 3376 (35.7 x 95.0) 1. Frey scope Lake Mich – sta, l

Slide from ANSP 46906a May 1947 L. Mich. Chicago (42.2 x 119.2) y

FIGURE 2. Notes from reverse of card shown in Figure 1. Because *Amphora calumetica* is relatively rare, emphasis is on locating a range of specimens.



Amphora calumetica (43.4 x 117.2) y
85.5/11
spec.
NEG. 60 G
Photo mag. 1000 X Coll.# 1574

FIGURE 3. Example of an ancillary card, showing largest specimen of *A. calumetica* found at the time.

the card (Fig. 2). More than one card could be used to illustrate morphological variation and size series of any given entity (Fig. 3). Of course this is all very primitive, given the current availability of excellent databases that easily incorporate such information and are very easy to use. An example is the Filemaker™ template developed by Joynt (Joynt and Wolfe 1999) that can incorporate all these features and considerably more. The really important aspect of using such a system, rather than relying entirely on the published literature is that it allow one to follow the dictum of "when in doubt, sort it out." In the case of the Great Lakes, it was obvious that many "common species" had different morphotypes that had separate distribution patterns (Pappas and Stoermer 2001), and likely were genetically separate entities. Although separation of taxa on minor morphological variations might seem risky, in terms of supporting ecological interpretation, it is vastly less destructive than under-classification (Birks 1994). In fact, most multivariate statistical techniques will, given that identification is consistent, merely re-aggregate false separations.

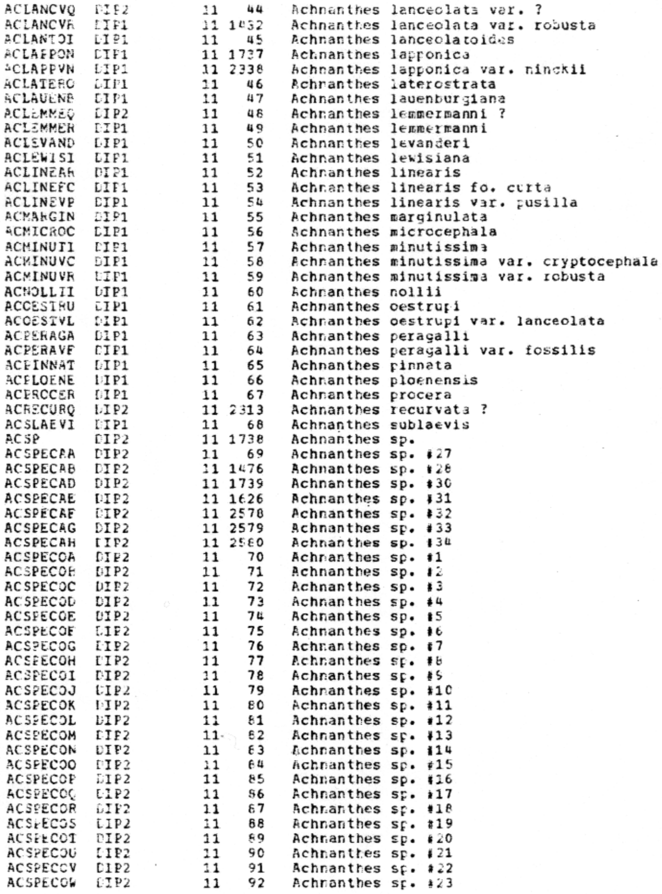
Computerization

In our case, computerization began as a simple data analysis problem. When handling large data sets, verification and data integrity are always problems, and ones that humans seem to handle poorly. Remembering these were the days when computer memories were limited and storage devices primitive. We had quite a struggle with programmers to use names recognizable to humans, and let the computer do the lookup, rather than simplifying the programmer's task by using a simple sequential list of taxa. Although this seems trivial in the modern context, I think there is an important lesson. Let computers do the simple, purely logical tasks. Save the human ability to deal with more complex tasks, perhaps aided by calculating engines, for the hard parts.

From this humble beginning, we, largely through the efforts of Theodore and Barbara Ladewski, were able to develop in integrated database system useful to both taxonomy and ecology. The program's name, through its several incarnations, was FIDO (a programmer's play on the word "phyto"). It consisted of the following elements:

MASTERLIST — A list of all acceptable names. These could be in the form of proper Latin binomials, binomials of convenience, or simple numerical or other arbitrary designation. The important part was that in order to become part of Masterlist, any designation had to be supported by a marked specimen in the collection and a photographic illustration in the master card file. Of course, all of these functions can be incorporated in any modern database. An sample portion is shown in Figure 4.

DECKCHECK — a subprogram that checked all entries for codes not acceptable to Masterlist (coding violations, misspellings, etc.) and "suspicious" data. I am surprised at how few current databases include extended data verification protocols. It is our experience that an appreciable error rate is associated with human data entry and review, no matter how careful the analyst or transcriber, and many of these can be detected by fairly simple data screening protocols.



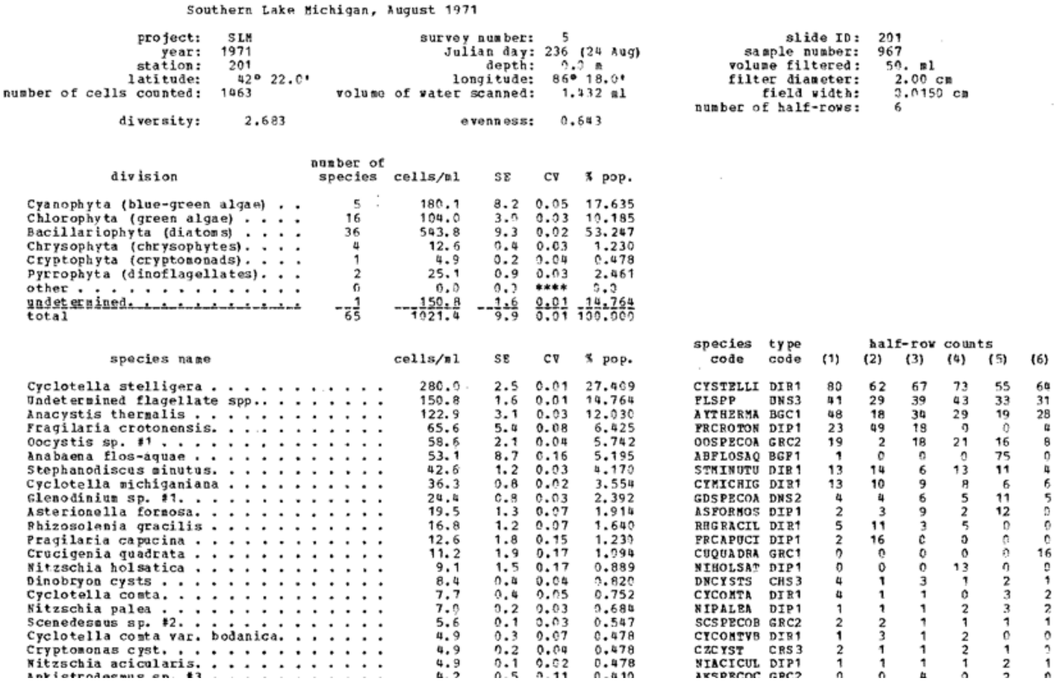
```
ACLANCVQ  DIP2       11   44  Achnanthes lanceolata var. ?
ACLANCVR  DIP1       11 1432  Achnanthes lanceolata var. robusta
ACLANTOI  DIP1       11   45  Achnanthes lanceolatoides
ACLAPPON  DIP1       11 1737  Achnanthes lapponica
ACLAPPVN  DIP1       11 2338  Achnanthes lapponica var. ninckii
ACLATERO  DIP1       11   46  Achnanthes laterostrata
ACLAUENB  DIP1       11   47  Achnanthes lauenburgiana
ACLEMMEQ  DIP2       11   48  Achnanthes lemmermanni ?
ACLEMMER  DIP1       11   49  Achnanthes lemmermanni
ACLEVAND  DIP1       11   50  Achnanthes levanderi
ACLEWISI  DIP1       11   51  Achnanthes lewisiana
ACLINEAR  DIP1       11   52  Achnanthes linearis
ACLINEFC  DIP1       11   53  Achnanthes linearis fo. curta
ACLINEVP  DIP1       11   54  Achnanthes linearis var. pusilla
ACMARGIN  DIP1       11   55  Achnanthes marginulata
ACMICROC  DIP1       11   56  Achnanthes microcephala
ACMINUTI  DIP1       11   57  Achnanthes minutissima
ACMINUVC  DIP1       11   58  Achnanthes minutissima var. cryptocephala
ACMINUVR  DIP1       11   59  Achnanthes minutissima var. robusta
ACNOLLII  DIP1       11   60  Achnanthes nollii
ACOESTRU  DIP1       11   61  Achnanthes oestrupi
ACOESTVL  DIP1       11   62  Achnanthes oestrupi var. lanceolata
ACPERAGA  DIP1       11   63  Achnanthes peragalli
ACPERAVF  DIP1       11   64  Achnanthes peragalli var. fossilis
ACPINNAT  DIP1       11   65  Achnanthes pinnata
ACPLOENE  DIP1       11   66  Achnanthes ploenensis
ACPROCER  DIP1       11   67  Achnanthes procera
ACRECURQ  DIP2       11 2313  Achnanthes recurvata ?
ACSLAEVI  DIP1       11   68  Achnanthes sublaevis
ACSP      DIP2       11 1738  Achnanthes sp.
ACSPECAA  DIP2       11   69  Achnanthes sp. #27
ACSPECAB  DIP2       11 1476  Achnanthes sp. #28
ACSPECAD  DIP2       11 1739  Achnanthes sp. #30
ACSPECAE  DIP2       11 1626  Achnanthes sp. #31
ACSPECAF  DIP2       11 2578  Achnanthes sp. #32
ACSPECAG  DIP2       11 2579  Achnanthes sp. #33
ACSPECAH  DIP2       11 2580  Achnanthes sp. #34
ACSPECOA  DIP2       11   70  Achnanthes sp. #1
ACSPECOB  DIP2       11   71  Achnanthes sp. #2
ACSPECOC  DIP2       11   72  Achnanthes sp. #3
ACSPECOD  DIP2       11   73  Achnanthes sp. #4
ACSPECOE  DIP2       11   74  Achnanthes sp. #5
ACSPECOF  DIP2       11   75  Achnanthes sp. #6
ACSPECOG  DIP2       11   76  Achnanthes sp. #7
ACSPECOH  DIP2       11   77  Achnanthes sp. #8
ACSPECOI  DIP2       11   78  Achnanthes sp. #9
ACSPECOJ  DIP2       11   79  Achnanthes sp. #10
ACSPECOK  DIP2       11   80  Achnanthes sp. #11
ACSPECOL  DIP2       11   81  Achnanthes sp. #12
ACSPECOM  DIP2       11   82  Achnanthes sp. #13
ACSPECON  DIP2       11   83  Achnanthes sp. #14
ACSPECOO  DIP2       11   84  Achnanthes sp. #15
ACSPECOP  DIP2       11   85  Achnanthes sp. #16
ACSPECOQ  DIP2       11   86  Achnanthes sp. #17
ACSPECOR  DIP2       11   87  Achnanthes sp. #18
ACSPECOS  DIP2       11   88  Achnanthes sp. #19
ACSPECOT  DIP2       11   89  Achnanthes sp. #20
ACSPECOU  DIP2       11   90  Achnanthes sp. #21
ACSPECOV  DIP2       11   91  Achnanthes sp. #22
ACSPECOW  DIP2       11   92  Achnanthes sp. #23
```

FIGURE 4. A fragment of MASTERLIST printed in the late 1970s. Reading from the left, identity code, a major group and habitat code, two columns of numerical book keeping codes used by the program, and accepted epithets. At present, only about 20% of arbitrary numerical designations have been identified with described species,

TAPEIT — A subprogram that wrote files for further processing and a separate permanent archive.

FETCH — A subprogram that retrieved data from the archive, either as hardcopy with summary statistics (subprogram ANALYZE) or output for further manipulation. An example of the former is shown in Figure 5. Note that summary statistics are calculated, including error estimates on counts. A separate, parallel-running system was used to collect and process chemical and physical data. This system was structured similarly to FIDO, which made merging of the databases for analysis relatively simple (Fig. 6). Examples of further manipulations include such things as distribution maps (Fig. 7) and representations of community structure based on multivariate statistical analyses (Figs. 8 and 9).

In the discussion above readers will note that almost all the design criteria were motivated by trying to bring some sort of modern taxonomic understanding to relatively large scale ecological



```
                        Southern Lake Michigan, August 1971

              project:  SLM                  survey number:  5                          slide ID:  201
                 year:  1971                   Julian day: 236 (24 Aug)            sample number:  967
              station:  201                         depth:  0.0 m               volume filtered:   50. ml
             latitude:   42° 22.0'              longitude:  86° 18.0'           filter diameter:    2.00 cm
number of cells counted:  1063        volume of water scanned:  1.432 ml            field width:    0.0150 cm
                                                                          number of half-rows:    6
            diversity:     2.683                    evenness:   0.643

                                   number of
            division                species   cells/ml    SE     CV    % pop.

   Cyanophyta (blue-green algae) . .     5       180.1    8.2   0.05   17.635
   Chlorophyta (green algae) . . . .    16       104.0    3.0   0.03   10.185
   Bacillariophyta (diatoms) . . . .    36       543.8    9.3   0.02   53.247
   Chrysophyta (chrysophytes). . . .     4        12.6    0.4   0.03    1.230
   Cryptophyta (cryptomonads). . . .     1         4.9    0.2   0.04    0.478
   Pyrrophyta (dinoflagellates). . .     2        25.1    0.9   0.03    2.461
   other . . . . . . . . . . . . . .     0         0.0    0.0   ****    0.0
   undetermined. . . . . . . . . . .     1       150.8    1.6   0.01   14.764
   total                                65      1021.4    9.9   0.01  100.000
                                                                                species  type             half-row counts
            species name                      cells/ml    SE     CV    % pop.    code    code   (1)   (2)   (3)   (4)   (5)   (6)

   Cyclotella stelligera . . . . . . . . . .   280.0    2.5   0.01   27.409    CYSTELLI DIR1   80    62    67    73    55    64
   Undetermined flagellate spp.. . . . . . .   150.8    1.6   0.01   14.764    FLSPP    DNS3   41    29    39    43    33    31
   Anacystis thermalis . . . . . . . . . . .   122.9    3.1   0.03   12.030    AYTHERMA BGC1   48    18    34    29    19    28
   Fragilaria crotonensis. . . . . . . . . .    65.6    5.4   0.08    6.425    FRCROTON DIP1   23    49    18     0     0     4
   Oocystis sp. #1 . . . . . . . . . . . . .    58.6    2.1   0.04    5.742    OOSPECOA GRC2   19     2    18    21    16     8
   Anabaena flos-aquae . . . . . . . . . . .    53.1    8.7   0.16    5.195    ABFLOSAQ BGF1    1     0     0     0    75     0
   Stephanodiscus minutus. . . . . . . . . .    42.6    1.2   0.03    4.170    STMINUTU DIR1   13    14     6    13    11     4
   Cyclotella michiganiana . . . . . . . . .    36.3    0.8   0.02    3.554    CYMICHIG DIR1   13    10     9     8     6     6
   Glenodinium sp. #1. . . . . . . . . . . .    24.4    0.8   0.03    2.392    GDSPECOA DNS2    4     4     6     5    11     5
   Asterionella formosa. . . . . . . . . . .    19.5    1.3   0.07    1.914    ASFORMOS DIP1    2     3     9     2    12     0
   Rhizosolenia gracilis . . . . . . . . . .    16.8    1.2   0.07    1.640    RHGRACIL DIR1    5    11     3     5     0     0
   Fragilaria capucina . . . . . . . . . . .    12.6    1.8   0.15    1.230    FRCAPUCI DIP1    2    16     0     0     0     0
   Crucigenia quadrata . . . . . . . . . . .    11.2    1.9   0.17    1.094    CUQUADRA GRC1    0     0     0     0     0    16
   Nitzschia holsatica . . . . . . . . . . .     9.1    1.5   0.17    0.889    NIHOLSAT DIP1    0     0     0    13     0     0
   Dinobryon cysts . . . . . . . . . . . . .     8.4    0.4   0.04    0.820    DNCYSTS  CHS3    4     1     3     1     2     1
   Cyclotella comta. . . . . . . . . . . . .     7.7    0.4   0.05    0.752    CYCOMTA  DIR1    4     1     1     0     3     2
   Nitzschia palea . . . . . . . . . . . . .     7.0    0.2   0.03    0.684    NIPALEA  DIP1    1     1     1     2     3     2
   Scenedesmus sp. #2. . . . . . . . . . . .     5.6    0.1   0.03    0.547    SCSPECOB GRC2    2     2     1     1     1     1
   Cyclotella comta var. bodanica. . . . . .     4.9    0.3   0.07    0.478    CYCOMTVB DIR1    1     3     1     2     0     0
   Cryptomonas cyst. . . . . . . . . . . . .     4.9    0.2   0.04    0.478    CZCYST   CRS3    2     1     1     2     1     0
   Nitzschia acicularis. . . . . . . . . . .     4.9    0.1   0.02    0.478    NIACICUL DIP1    1     1     1     1     2     1
   Ankistrodesmus sp. #3 . . . . . . . . . .     4.2    0.5   0.11    0.410    AKSPECOC GRC2    0     0     4     0     2     0
```

FIGURE 5. Example of ANALYZE output taken from a study of whole phytoplankton (diatoms and other groups) in southern Lake Michigan in 1971. Raw data are shown in right hand columns. Summarized data are shown in left columns. The large "undetermined" category consists mostly of microflagellates that cannot be satisfactorily identified with light microscopy.

projects, lacking the sort of traditional floristic and monographic support generally assumed. Perhaps more importantly, once our national science funding establishment began to awake to the fact that we are living in an ecosystem that is probably less than 20% described, this type of data base made it possible to attack some real taxonomic problems, particularly of the Great Lakes region (e.g., Theriot and Stoermer 1984, 1986).

MISTAKES AND PROBLEMS

In retrospect, it is nearly always possible to identify mistaken directions and things that should have been done differently. In our case the worst problems were partially our own fault and partially due to faults in the system. Part of the problem was that we started early in the game. Many diatomists resisted computer applications when they first became available. On the other hand, the funding agencies we dealt

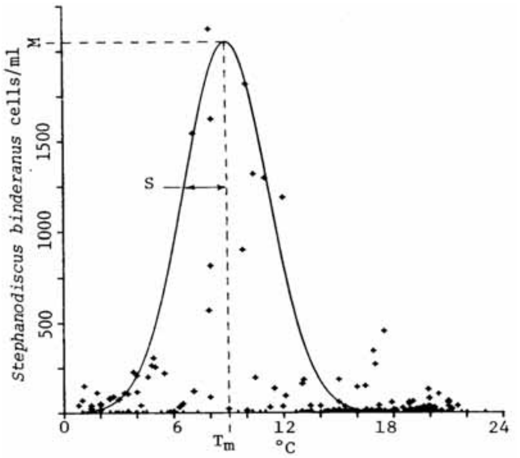


FIGURE 6. Example of data plotted from the study cited in Figure 5, in this case the absolute abundance of *Stephanodiscus binderanus* (Kütz.) Krieg relative to temperature (Stoermer and Ladewski 1976). Curve is fitted to data envelope and estimates of maximum abundance (M) and dispersion (S) are derived. Anomalous appearing points on the right come from inshore stations in the fall when populations are injected into the still warm lake from more rapidly cooling streams.

with at the time were reluctant to provide support dedicated to database development at the local project level. Some spent inordinate amounts of money on commercial database development, but most of these were put together with minimal inputs from the user community and, although they might have incorporated the latest programming tricks of the time, were hideously clumsy and inefficient to use. It has been my observation that most really useful databases incorporate a high level of specific user input, and most really successful programs are locally developed. Since computerization has become popular our national funding agencies have devoted considerable resources to development of several generations of biological databases, but most of this effort has gone to generalized systems that are not particularly appropriate for the problems faced by diatomists.

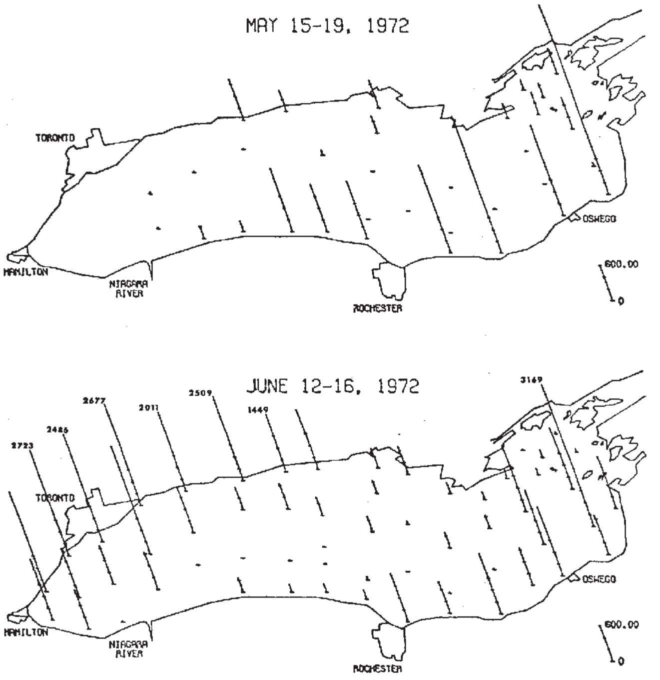


FIGURE 7. Example of species data plotted from a similar study. Distribution of *S. binderanus* in Lake Ontario in the spring of 1972. In the lower image, actual numerical values are given at the top of bars when values are too large to conveniently plot at scale used (from Stoermer et al. 1974).

Part of the problem is the structure and economics of the computer industry. The very rapid expansion of computing power (Moore's Law) causes rapid obsolescence in microcomputers, a trend that the industry has capitalized on. It must also be said that University administrations, at least in this country, have been alert to the fact that the cost of centralized mainframe computer systems usually becomes their responsibility, whereas much of the cost of decentralized systems falls on Departments, or individual investigators. It is also a truism that the quickest way for a software company to go broke is to design a perfect product. It is economically much more rewarding to design something marginally adequate that can continue to be upgraded. All of this militates against development of a stable continuing system, and makes upgrading of a developed system very difficult, in that most resources are devoted to exploiting "exciting" new technologies, rather than adapting existing databases to them as they arise.

In the case of our system described above, we eventually became victims of the technology transition. FIDO was much more complete and easy to use than any of the early microcomputer database programs, and we continued to use it well past the transition from mainframe-based to a microcomputer-based network system. We were unable to obtain support for conversion from either local or national funding sources, so much of the data accumulated during this era exists only on hardcopy and tapes that are rapidly becoming unreadable. Part of the reason for this was that we were somewhat too clever in using "latest technologies" of the day that were specific to the University of Michigan mainframe computer system.

Perhaps the "take home" message for independent laboratories is to develop and use the simplest system adequate to your specific needs, and upgrade and maintain it judiciously. Although the latest and greatest in technology is always attractive, pioneers in technology areas often suffer different, but equally painful, slings and arrows as did the geographic pioneers of past centuries. In this regard, I think the "open software" movement offers great promise.

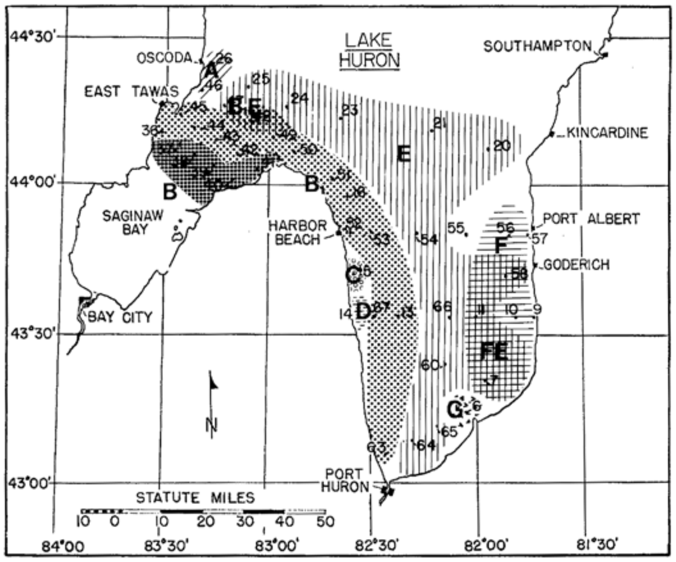


FIGURE 8. Representation of phytoplankton community structure in southern Lake Huron based on samples taken 4-8 June 1974 under west wind forcing. Associations were determined using dimensional ordination and principal components analysis (from Stoermer and Kreis 1980). Materials and phytoplankton from badly polluted Saginaw Bay are entrained by the spring thermal bar and, combined with other local shoreline sources, generate "eutrophic" associations in the western portion of the lake. Mostly agricultural and minor industrial sources from the Canadian shore, also entrained by the spring thermal bar, produce more "mesotrophic" associations in the eastern portion of the lake. The oligotrophic associations expected in a large lake of this type are only found in the offshore waters.

PRESENT AND FUTURE CONSIDERATIONS

And I continue to feel that computer assisted approaches offer the best avenue for "marrying" the needs of taxonomists and ecologists. As I have discussed elsewhere (Stoermer 2001) it is foolish for ecologists to expect taxonomic treatises on diatoms of the type generally available for "higher" organisms to become available in the foreseeable future. This being the case, it is really necessary to incorporate good taxonomic practice into routine analytical work and assure that project outputs are useful to people whose primary interests are in taxonomy and systematics. At the same time, it behooves the few people in the latter category to be more proactive in addressing the resources potentially available from ecological studies.

At present, it is quite feasible for workstations used in diatom analysis to capture and maintain not only the analysts' taxonomic decisions, but also images of exemplar specimens such decisions are based on, the pertinent locality information, and the precise location on a slide of each specimen assigned to a given category. At the same time, the analyst should be able to address taxonomic information and identification aids, such as image analysis, directly and in real time.

Whereas the digital tools now available offer exciting possibilities, they also present some real challenges and dangers. The possibilities for enhanced data display and sharing make the possibility of "consensus floras" more attractive. Although this may be useful, and indeed necessary, in the context of a particular ecological project, such efforts can easily degenerate into lowest common denominator solutions that actually retard scientific progress in the general field, rather than advancing it. Diatomists are in a particularly difficult situation in this regard. Taxonomic information in our field is virtually exploding, but most funding agencies, both those traditionally supporting ecological research and those supporting taxonomic tend to take large organisms as their model for understanding diversity. Even at this level, there is no logical expectation of ever establishing a truly "stable" taxonomic system unless we are willing to freeze knowledge in some imperfect state. In the case of diatoms, the present state is grossly imperfect and the expectation of stability is demonstrably unscientific. Given that there are snares and pitfalls to be avoided, currently available technologies offer those bold and resourceful enough to utilize them great possibilities. These range from purely exploratory — we are still in the era where simple discovery and description probably advances the field more than any other approach — to application and incorporation of available tools for taxonomic and ecological questions.

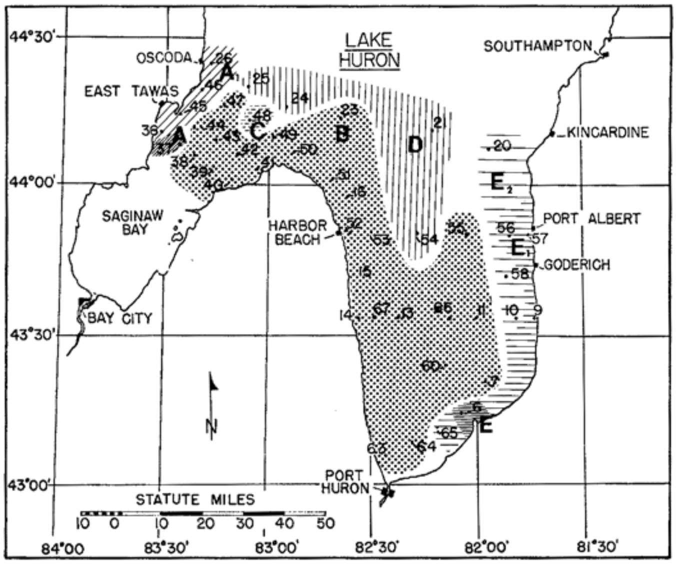


FIGURE 9. Representation of phytoplankton associations from the same study shown in Figure 8. In this case, data were collected 26-31 August under east wind forcing. A large upwelling has occurred in the eastern region of the lake. This combined with local shoreline sources results in atypical phytoplankton associations in the eastern nearshore region. The extent of nutrient re-supply also causes somewhat atypical summer associations in most of the southern portion of the lake, and these communities intrude into Saginaw Bay, as the expected eutrophic communities are transported northward along the Michigan (western) shore. The expected offshore "oligotrophic" summer phytoplankton association is only found at a few stations in the north-central quarter.

REFERENCES

BAILEY, J.W. 1842a. A sketch of the Infusoria of the Family Bacillaria, with some accounts of the most interesting species which have been found in a recent or fossil state in the United States. *Ameican Journal of Science and Arts* 42: 88–105.

BAILEY , J.W. 1842b. A sketch of the Infusoria of the Family Bacillaria, with some accounts of the most interesting species which have been found in a recent or fossil state in the United States. *Ameican Journal of Science and Arts* 42: 321–332.

BEATTY, W.K. 1982. When Cholera Scourged Chicago. *Chicago History* 11:2–13.

BEETON, A.M. 1965. Eutrophication of the St. Lawrence Great Lakes. *Limnology and Oceanography* 10:240–254.

BEETON, A.M. 1969. Changes in the environment and biota of the Great Lakes. Pages 150–187 in *Eutrophication: Causes, Consequences, Correctives*. National Academy of Sciences, Washington, DC, USA.

BIERMAN, V.J., JR., AND D.M. DOLAN. 1981. Modeling of phytoplankton-nutrient dynamics in Saginaw Bay, Lake Huron. *Journal of Great Lakes Research* 7:409–439.

BIRKS, H.J.B. 1994. The importance of pollen and diatom taxonomic precision in quantitative paleoenvironmental reconstructions. *Revue of Paleobotany and Palynology* 83: 107–117.

BONNER, T.M. 1991. *Medicine in Chicago 1850-1950*, 2nd ed. University of Illinois, Urbana, Illinois, USA.

EHRENBERG, C.G. 1845. Neue Untersuchungen über das kleinste Leben als geologisches Moment. *Berichte über die zur bekanntmachung geeigneten Verhandlungen der Koniglich-Preussischen Akademie der Wissenschaften zu Berlin* 1845:53-88.

EHRENBERG, C.G. 1854. *Zur Mikrogologie: Das Erden und Felsen schaffende Wirken des unsichtbar selbst-*

ständigen Lebens auf der Erde. Leopold Voss, Leipzig. 374 pp. + Fortsetzung (1856). 88 pp.

JOHNSTON, E.M., AND E.F. STOERMER. 1976. Computer analysis of phytoplankton cell images. *The Microscope* 248: 181–187.

JOYNT E.H., III, AND A.P. WOLFE. 1999. An image database for diatom identification and nomenclature. *Journal of Paleolimnology* 22:109–114.

LADEWSKI, T.B., AND E. F. STOERMER. 1973. Water transparency in southern Lake Michigan in 1971 and 1972. Pages 791–807 in *Proceedings of the 16th Conference on Great Lakes Research*. International Association for Great Lakes Research.

PAPPAS, J.L., AND E.F. STOERMER. 2001. Fourier shape analysis and fuzzy measure shape group differentiation of Great Lakes *Asterionella* Hassall (Heterokontophyta, Bacillariophyceae). Pages 485–501 *in* A. Economou-Amilli, ed., *Proceedings of the Sixteenth International Diatom Symposium, Athens*. Amvrosiou Press, Athens, Greece.

SICKO-GOAD, L., E.F. STOERMER, AND B.G. LADEWSKI. 1977. A morphometric method for correcting phytoplankton cell volume estimates. *Protoplasma* 93:147–163.

SMITH, S.H. 1972. Destruction of the ecosystem in the Great Lakes and possibilities for its reconstruction. *University of Washington Publications in Fisheries*, New Series 5:41–46.

STOERMER, E.F. 2001. Diatom taxonomy for paleolimnologists. *Journal of Paleolimnology* 25:393–398.

STOERMER, E.F., M.M. BOWMAN, J.C. KINGSTON, AND A.L. SCHAEDEL. 1974. *Phytoplankton Composition and Abundance in Lake Ontario during IFYGL*. University of Michigan, Great Lakes Research Division, Special Report No. 53. 373 pp.

STOERMER, E.F., AND R.G. KREIS, JR. 1980. *Phytoplankton Composition and Abundance in Southern Lake Huron*. U.S. Environmental Protection Agency, Research Reporting Series EPA-600/3-80-061, Environmental Research Laboratory, Duluth, Minnesota, USA. 382 pp.

STOERMER, E.F., AND T.B. LADEWSKI. 1976. *Apparent Optimal Temperatures for the Occurrence of Some Common Phytoplankton Species in Southern Lake Michigan*. University of Michigan, Great Lakes Research Division, Publication No. 18. 49 pp.

STOERMER, E.F., AND R.J. STEVENSON. 1980. *Green Bay phytoplankton composition, abundance and distribution*. U.S. Environmental Protection Agency, Great Lakes National Program Office, Chicago, Ill. 103 pp.

STOERMER, E.F., AND T.B. LADEWSKI. 1982. Quantitative analysis of shape variation in type and modern populations of *Gomphoneis herculeana*. *Nova Hedwigia*, *Beih*. 73:347–386.

STOERMER, E.F., AND E.C. THERIOT. 1985. Phytoplankton distribution in Saginaw Bay. *Journal of Great Lakes Research* 11:132–142.

THERIOT, E.C., AND E.F. STOERMER. 1984. Principal component analysis of *Stephanodiscus*: Observations on two new species from the *Stephanodiscus niagarae* complex. *Bacillaria* 7:37–58.

THERIOT, E.C., AND E.F. STOERMER. 1986. Principal components analysis of *Stephanodiscus*: Field evidence for two varieties of *S. niagarae*. Pages 385–394 in *Proceedings of the VIIIth International Symposium on Recent and Fossil Diatoms, Paris*. Koeltz Scientific Books, Koenigstein, Germany.

THOMAS, B.W., AND H.H. CHASE. 1887. Diatomaceae of Lake Michigan as collected during the last 16 years from the water supply of the City of Chicago. *Notarisia* 2:328–330.

THOMPSON, H.D. 1896. Report on the plants. *Bulletin of the Michigan Fisheries Commission* 6:72–75.

VAUGHN, J.C. 1961. Coagulation difficulties of the south district filtration plant. *Pure Water* 13:45–49.

VAUGHN, J.C. 1962. Operation of the south district filtration plant, 1961. *Pure Water* 14:34–49.

VORCE, C.M. 1881. Forms observed in water of Lake Erie. *Transactions of the American Microscopical Society* 4:51–90.

VORCE, C.M. 1882. Microscopic forms observed in the waters of Lake Erie. *Proceedings of the American Society of Microscopy* 5:187–196.

WARD, H.H. 1896. A biological examination of Lake Michigan in the Traverse Bay Region. *Bulletin of the Michigan Fisheries Commission* 6:1–71.

WOLFE, T., AND R. SWEENEY. 1980. *Laurentian Great Lakes* Cladophora*: Annotated Bibliography*. Great Lakes Laboratory, State University College at Buffalo, Buffalo, New York, USA. 54 pp.