Field Station 2000 Initiative

Rationale and action plan for networking the Organization of Biological Field Stations (OBFS) to empower demonstration of national environmental conditions and trends.

http:www.obfs.org

Results of a Workshop

May 17-22, 1998

at
The National Center for Ecosystem Analysis and Synthesis
Santa Barbara, California

Revised per discussion and the Annual Meeting at Archbold Field Station,
September, 1998

February 24, 1999

Report of the OBFS Working Group on Networking

Stanford, Jack - Co-Chair
Flathead Lake Biological Station
stanford@selway.umt.edu

McKee, Art – Co-Chair
Andrews Experimental Forest
mckee@fsl.orst.edu

Allison, Taber
Rocky Mountain Biological Lab
tabermbl@crestedbutte.net

Brunt, James
LTER Office, University of New Mexico
jbrunt@lternet.edu

Callahan, Tom
National Science Foundation
jcallaha@nsf.gov

Cohen, Philippe
Jasper Ridge Natural Area
philippe@jasper.stanford.edu

Dalgleish, Bob
University of Missouri – St. Louis
dalgleish@umsl.edu

Helly, John
San Diego Supercomputer Center
Hellyj@ucsd.edu

Holland, Marge
University of Mississippi Field Station
mholland@olemiss.edu

Kennedy, Jeff
University of California, Davis
jakennedy@ucdavis.edu

Michener, William
Jones Ecological Research Center
wmichene@jonesctr.org

Stromberg, Mark
Hastings Reserve
stromber@socrates.berkeley.edu

Swain, Hilary
Archbold Biological Station
hswain@archbold-station.org

White, David
Hancock Biological Station
david.white@murraystate.edu

Wyman, Richard
E.N. Huyck Preserve
rlwyman@capital.net

EXECUTIVE SUMMARY

- Representatives of the Organization of Biological Field Stations (OBFS) met at the National Center for Ecological Analysis and Synthesis (NCEAS) to examine rationale and impediments for shared data collection and management by a national network of OBFS members. The premise is that modern electronic communications can facilitate a much higher resolution of strategic environmental issues through the distributed information system represented by OBFS.

- Currently, environmental data routinely collected at the 160 OBFS sites in North America are not systematically archived in electronic media. Therefore, the massive data and other information resources collectively held by OBFS sites are not easily accessible for analysis and synthesis of strategic environmental issues and questions.

- The working group developed a framework for the establishment of an Internet-based network for data sharing and archiving among OBFS sites. The framework involves OBFS member stations, Long Term Ecological Research (LTER) sites (most of which are OBFS members), the National Center for Ecological Analysis and Synthesis (NCEAS) and the San Diego Supercomputing Center (SDSC) in a consortium for coordinated data collection, analysis, synthesis, visualization and electronic archival.

- Three projects to illustrate the value of such a network were developed. First, we used the existing LTER protocol to establish a personnel data base for all OBFS member stations. The protocol has been set up on the OBFS webpage (http://jasper.stanford.edu/OBFS/ index.html) and OBFS stations currently are adding their personnel information to this data base. OBFS member stations should immediately update their personnel directory per instructions on the OBFS web site.

- A second project demonstrated the scientific and monitoring utility of a network by surveying the status of amphibian populations at a suite of field stations. The results of the survey may be viewed on the OBFS webpage and are summarized herein. In a third project, which we have not yet begun, we will compile species lists from all stations as the first step in a potentially novel analysis of regional and national patterns of biodiversity. We anticipate that these three projects taken together will demonstrate the feasibility of advanced analyses using distributed data available at OBFS sites.

- Protocols and standards for data management and sharing through the development of an OBFS network are proposed and discussed herein, including training workshops, and development of support teams and tools. Our plans are derived in large measure from a recent OBFS-LTER-ESA workshop that produced a document entitled Data and Information Management in the Ecological Sciences: A Resource Guide (DIMES report, Michener et al. 1998).

- An OBFS network will provide the information and many of the analyses necessary to scientifically answer vitally important questions about the current and future condition of strategic natural resources nationwide. Strategic resources include clean and sustained sources of air, water, food, fiber and other products provided by terrestrial, freshwater and marine ecosystems. Hence, we will contribute to the development of more informed land use and environmental policies at local to national levels.

- Access to and participation in a national network for environmental research and monitoring also will enhance education, which is a primary mission at most OBFS sites. The OBFS network will facilitate proactive involvement of students and faculty in collecting, analyzing and synthesizing shared data bases to resolve national environmental quality issues.

- Future tasks

    1. *Staff an OBFS network office with a website and distributed information manager.* The immediate objectives are: 1) to develop a directory of OBFS sites and personnel data base, using the LTER protocol, 2) develop an OBFS information locator system and, 3) to inventory and catalog long-term environmental data bases that currently are available at the 158 OBFS sites. If possible, the OBFS network office may be in collaboration with LTER and hopefully involve NSF funding.
    2. *Continue development of data acquisition and management initiatives collaboratively with the LTER, SDSC and other organizations and agencies*, particularly AERC, NAML, NASA, EPA and USGS (all have objectives related to the OBFS network goal). We will approach field sites nationally that are not now members of OBFS (e.g., certain Nature Conservancy Sites, U.S. Forest Service Experiment Stations) for possible inclusion in the OBFS network.
    3. *Investigate the feasibility of expanding the OBFS network to International Organization of Biological Field Stations (IOBFS).*
    4. *Conduct a workshop sometime in late 1999 or early 2000 to develop an implementation and funding plan for enlarging the scope and effectiveness of the OBFS network for national and global environmental monitoring (field stations as earth observatories).* Key objectives include: 1) determination of a suite of key environmental variables and measurement protocols that can be effectively and routinely monitored at OBFS sites for the purpose of assessing local, regional and national trends and patterns of environmental change; 2) determination of whether our existing environmental data bases, as we expect, do contain strategic information about regional, national and global environmental change; and, 3) development of a strategy for funding and implementing network-based environmental change research and education.

- All information concerning the OBFS Field Stations 2000 Initiative will be posted on the OBFS web site, including the full text of this report. The working group on networking (authors of this report) will continue to lead this effort to systematically network OBFS sites for national environmental monitoring and research.

I. INTRODUCTION

The nation is faced with uncertainty regarding the state and functioning of its ecological systems that cleanse water and air, facilitate production of food and fiber and sustain the many other natural resources upon which our quality of life depends. Increased ecological knowledge and understanding are needed to detect and monitor changes in ecosystems, to evaluate the consequences of human activities on ecosystem structure and function and to manage ecosystems and the services they provide in a sustainable manner. In response to these needs, the ecological scientific community has articulated research priorities (e.g., Lubchenco et al. 1991; Naiman et al. 1995) and proposed mechanisms for data storage, sharing and integration (e.g., Gross et al. 1995; Michener et al. 1998).

Biological field stations throughout the United States (Figure 1) represent a wealth of relevant ecological knowledge accumulated this century, but the research community often is unaware of these data sets or is unable to gain access to them. Moreover, an inventory of the massive set of environmental informatics (i.e., ecological data and information describing pattern and trends in climatic, hydrologic, chemical and biological variables) that biological field stations collectively hold has never been created and systematically updated in modern electronic formats. Very likely, these informatics are key resources for solving strategic environmental problems and for advancing basic science in ecology.

The New Horizons Report (Lohr et al. 1995), a collaborative effort of OBFS and the National Association of Marine Laboratories (NAML), lists key research themes for the next decade at biological field stations and marine labs:

- relating fundamentals of basic biology and ecology
- measuring environmental change
- maintaining biodiversity
- sustaining ecological systems
- predicting consequences of management policies and actions
- restoring and rehabilitating damaged ecosystems
- demonstrating rates of change in biological diversity and the subsequent effects on community structure and ecosystem processes
- describing the biology of rare and declining species and the scientific information necessary to sustain such species
- defining the principles that govern outbreak and spread of pest and disease organisms
- assessing evolutionary consequences of anthropogenic and other environmental changes.

This research agenda likely cannot be achieved at a national level of resolution without an effective framework for sharing and archiving informatics among biological field stations and marine labs.

To initiate its work, the group agreed that the primary elements of the informatics framework are:

1) a common understanding of strategic environmental questions that can be answered by collection, analysis and synthesis of OBFS data bases and other informatics (these are generally encompassed by the New Horizons research priorities and are not discussed further here, except with respect to mining new ideas and information from the existing inventory of data bases at OBFS sites and with respect to future workshops on the specific research questions or environmental issues);
2) a unified strategy for development, analysis and synthesis of regional and national data bases within a field station network to respond to these questions; and,
3) implementation of personnel and resources for collection, management, archiving and Internet posting of the required data bases at as many OBFS sites as practical and as defined by the respective missions of OBFS member stations.

The response of the working group to this informatics framework is presented in four sections that compose the body of this report. First (Section II), we propose a consortium for bioinformatics composed of OBFS member stations, the existing Long-Term Ecological Research network (all but 2 LTER sites are members of OBFS), the National Center for Ecological Analysis and Synthesis and San Diego Supercomputer Center. We then identify three Internet projects (site directory; amphibian survey; data base inventory) to demonstrate the utility of an OBFS network in research, education and environmental problem solving (Section III). In

Section IV, we examine information management as it relates to field station operations and development of a formal network. In the final sections of the report, we relate how we think the network can enhance the education mission of OBFS stations (Section V) and conclude with a formative plan for gradually expanding the scope and purpose of network (Section VI).

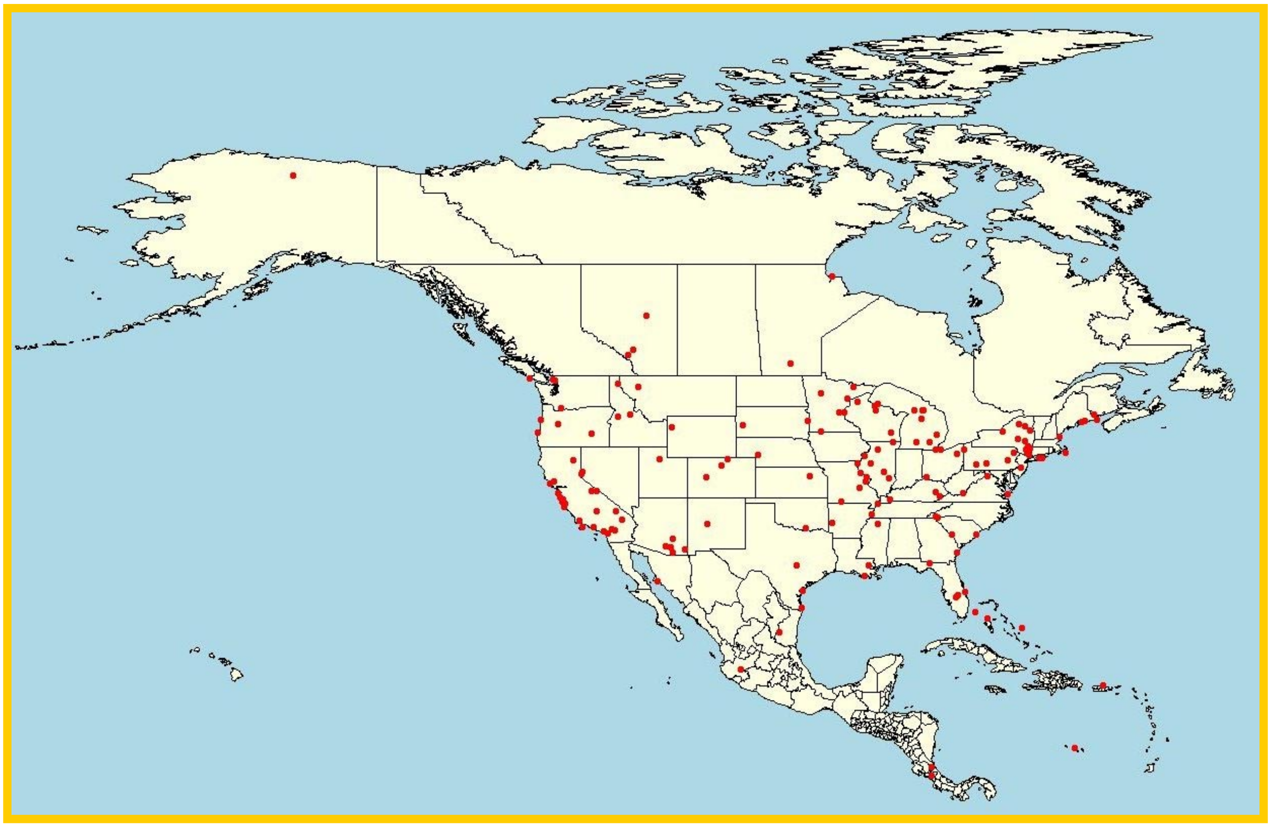

Figure 1. Locations of OBFS sites in North America.

II. PROPOSED CONSORTIUM

We envision proactive linkage of OBFS sites and objectives of the Long-Term Ecological Research (LTER) network for data base sharing and problem resolution. In the context of data management and sharing for the purpose of answering regional scale questions, LTER is the big brother of OBFS. However, the existing LTER network can be strengthened substantially by linkage to OBFS because of the number and strategic distribution of OBFS sites (Figure 1).

Collaboration of OBFS and LTER will be coordinated and enhanced on an as needed basis by the National Center for Ecological Analysis and Synthesis (NCEAS). The San Diego Supercomputer Center (SDSC) will be proactively involved in the OBFS-LTER network by functioning as the master data archiving center and assisting with implementation of advanced protocols for distributed information management, modeling and visualization. Collaboration with NCEAS and SDSC will insure that the OBFS-LTER network is developed in a state-of-the-art manner that will move to new computing, visualization, archiving and synthesis platforms in the most informed and efficient way possible.

The goal is to provide a master OBFS data base archive for analysis and synthesis of strategic environmental questions, as is currently being implemented for LTER sites and as is recommended in the Future Long-Term Ecological Data Report of the Ecological Society of America (Gross et al. 1995). We intend that the OBFS networking strategy compliment and eventually integrate with a similar effort by National Association of Marine Laboratories (NAML) referred to as LABNET. We also are hopeful that the Association of Ecosystem Research Centers (AERC) and the various professional societies in the ecological and environmental sciences will endorse and support this collaborative effort. And, as soon as possible, we envision expansion of the network to include the International Organization of Biological Field Stations (IOBFS).

III. NETWORK DEMONSTRATION PROJECTS

Three demonstrations at increasing levels of complexity are envisioned as a mechanism to demonstrate the utility of an OBFS network. The first, and probably least complex, is a 'personnel data base,' initially using information gathered from three field stations (Hastings National History Reserve, University of Mississippi Field Station and Archbold Biological Station). The second will use amphibian records from our field stations to demonstrate how the OBFS network might be used to answer a specific ecological question. The third is a systematic inventory of the collective data bases currently available at OBFS sites. In the near term, the latter effort will be limited to species lists and other data bases related to the general theme of biodiversity.

PERSONNEL DATA BASE

We have initiated development and implementation of a distributed information system encompassing all of the OBFS sites by utilizing the experience and infrastructure that has been developed by the LTER network. The LTER Network office has developed an Internet-accessible personnel data base. Each LTER site can access the system at SDSC for data entry over the Internet, using any browser software. With the help of the LTER network office we have utilized this system by OBFS as a starting point for the OBFS. This effort will be coordinated by Mark Stromberg (OBFS), John Helly (SDSC) and James Brunt (LTER).

A copy of the LTER personnel data base software was moved from Albuquerque to San Diego, where an expanded personnel data base will be maintained and include LTER and OBFS information. Each OBFS member station has a unique identifier and a unique web page for data entry. A common OBFS search page is being developed to allow any OBFS member to search the entire OBFS-LTER personnel data base. One of the most useful aspects of this scheme will be the capture of user data. Scientists and others working at OBFS stations can fill out individual data entries when they arrive at the stations. The entire data base can be easily searched.

Individual implementation of an OBFS-wide data base inevitably will require some modification at each site. For instance, some stations might want to use this system to classify users (faculty, graduate student, etc.) and capture data on which facilities are used and which buildings are used. Staff at each participating station will have to be identified and funded to implement and manage personnel at each station. As the system is implemented at more stations, the needs of individual stations to spend resources to join in this data network will decline. We will use this effort as a guideline for other more complex data sharing.

The OBFS personnel data base now resides at http://sql.lternet.edu/lterdb/OBFS.htm. The data base currently is populated only with information from Hastings (M. Stromberg, Director), University of Mississippi (M. Holland) and Archbold (H. Swain) Field Stations.

These stations exemplify the range of characteristics encompassing OBFS sites and, therefore, serve as examples for other stations to follow as they now add their data to the data base. We expect the data base to be complete by summer, 1999.

AMPHIBIAN DECLINE DEMONSTRATION PROJECT (ADDP)

The purpose of this project was to show how the OBFS network can rapidly assemble, edit and portray complex information using existing data and expertise within the OBFS sites (Figure 1). Our results are from this first of a three-phased effort to establish an environmental monitoring and research network among OBFS field sites. The ADDP is designed to demonstrate the potential synthetic value of a network among OBFS sites.

The work reported here demonstrates the feasibility of a national assessment of amphibian status and trends using the expertise typically found at biological field stations. Populations and species of amphibians appear to be declining nationally and some species have been locally extirpated while others may have recently become extinct (Wyman 1990). Additionally, recent reports of the increased frequency of deformities in several species of frogs and toads add urgency to understanding the current status of these organisms.

To accomplish this, we asked OBFS member stations to share information on amphibians and compile it on a national scale as a demonstration of the capacity of field stations to work together to give a national perspective on our biological heritage. We have successfully completed this demonstration and the results can be viewed on http://www.sdsc.edu/dx. Figure 2 is a representative plot of the data for one important genus using the web-site referenced above.

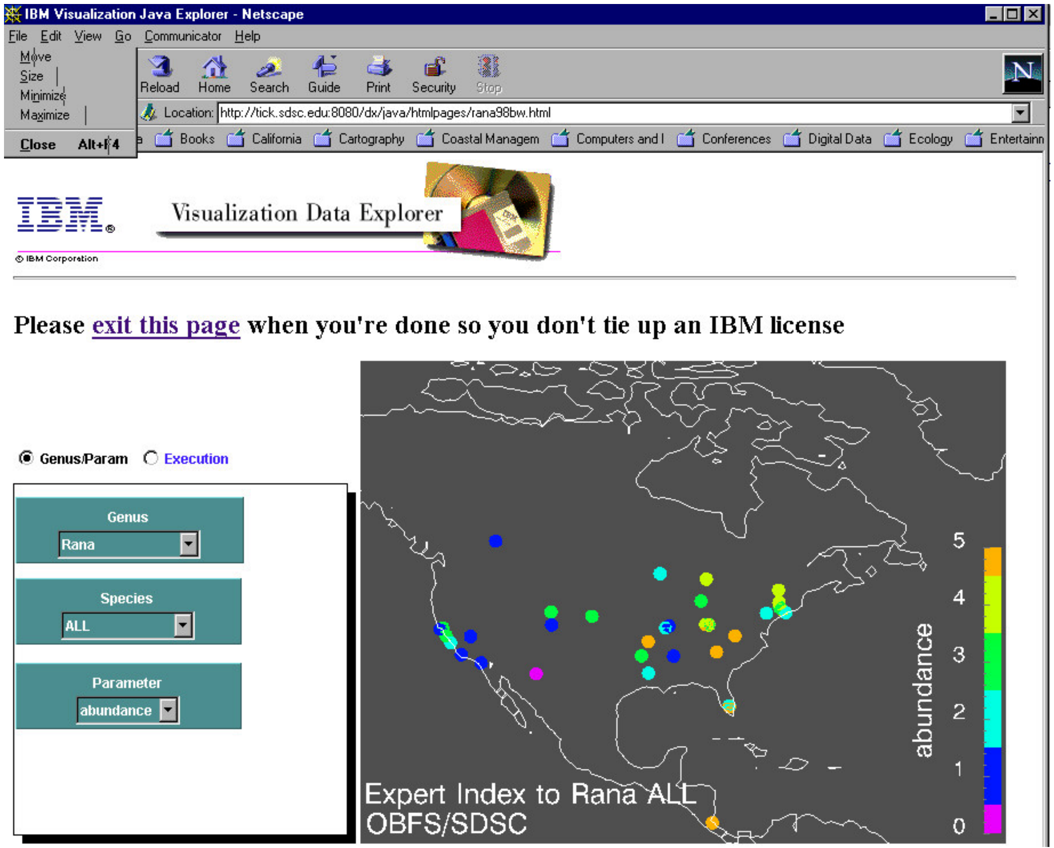

Figure 2. *Amphibian data display through a WWW browser using IBM Data Explorer Visualization Software with an application written by T. Todd Elvins at the San Diego Supercomputer Center. Data analysis and preparation was done by Mark Stromberg of UC Berkeley and John Helly of UC San Diego.*

We want to stress that these results should not be used alone to draw conclusions or to take regulatory actions. The data are presented here to demonstrate how a networked data collection and analysis effort, from existing biological field stations, can benefit the analysis of regional or national environmental problems. Data collected on amphibian populations vary tremendously in scope, methodology and detail. Recognizing the diversity of sampling methods used at OBFS sites in the absence of any prior attempt at coordination or standardization, we requested that the reporting scientists provide an informed opinion and expert evaluation of the abundance of all amphibians present in an area. Although, a quantitatively more rigorous data collection effort would require several years of planning and effort, Caughley (1977) points out that the majority of ecological problems can be tackled with the help of an index of abundance or density. Even non-linear indices are often adequate (Caughley, 1977) and are useful for comparisons. Thus, at the level of broad geographic comparisons (declining amphibians across North America) our indices are useful and appropriate for a pilot project.

We had a good response from most OBFS members. Only about 150 member stations are sufficiently mature to be expected to have the staff who have developed the expertise we requested. Those without the affiliated staff were often able to network with other biologists in the state or region who were interested in helping. Our response rate to the questionnaire (45 stations of 150) was thus very good. Through our pilot study collaboration we were able to rapidly handle incoming data and to provide the feedback loop so that widely distributed knowledge could be reviewed and corrected. Using the OBFS website and e-mail worked well for compiling and assuring data quality. Software and technical support at San Diego Supercomputer Center (SDSC) allow real time views of the data as it is compiled and corrected.

In general, data analyses like these enable meaningful comparisons across a large geographic scale. At this level, we can use the data to generate hypotheses which can lead to more detailed, local studies. A comparable use of indices of population data can be seen in a study where simple annual counts of acorn abundance across broad geographic scale lead to the discovery of large-scale synchrony in masting (Koenig and Knops, In press; Koenig et. al. 1998; Koenig and Knops, 1998). Other large-scale collection of indices to population data have been very useful, including the Audubon Christmas Bird Count or the USGS Breeding Bird Count (http://www.mp2-pwrc.usgs.gov/bbs/bbsops.htm). An exciting example of how a network of observers can compile useful data over the Internet has been Audubon's Great Backyard Bird Count: a one-day snapshot of 100 of the most common birds in the United States, involving over 14,000 observers (http://birdsource.cornell.edu/gbbc/index.html). Certainly, there are more people able to reliably recognize bird species than those trained to recognize amphibians. But the lessons learned in conducting such rapid, broad-scale acquisition and analysis of our biotic condition are clearly opening new avenues for helping policy makers and the public rapidly see and understand the desperate plight of our natural heritage.

INVENTORY OF OBFS DATA BASES

A primary objective of the NCEAS workshop that spawned this report was to develop a firm rationale and approach for demonstrating national trends in key environmental attributes. This idea derives from the wide distribution of stations across the USA (Figure 1) and the fact that almost all of these sites systematically gather environmental data as a basic part of their education and research missions. We believe that considerable information of national importance can be mined from OBFS data bases, but they need to be inventoried, quality controlled and made Internet accessible. The amphibian demonstration project, described above, strongly suggests that such information can be derived from a network of OBFS sites.

The breadth and status of data bases are expected to vary widely among OBFS member stations, as are current capacities for data base management. Moreover, the ability of stations to participate in a formal data base inventory will vary. For example, many, if not all OBFS stations, currently have species lists (birds, plants, fish, insects, etc.). A fewer number of stations would have abundance data to go with species lists, and even fewer would have species and abundance at a landscape scale. Nonetheless, this information, when formally assembled in a national data base, likely has enormous value in the context of a robust indicator of national environmental health. Moreover, many other types of data are collected at OBFS and LTER sites. The list below presents other types of data sets that may be available at OBFS sites. This is not meant to be exhaustive but illustrates potential areas of study.

- Spatial patterns
- Biological diversity
- Meteorology
- Hydrology
- Inorganic nutrients
- Carbon dynamics
- Paleoecology
- Secondary production

- Trophic dynamics

Based on our general knowledge of OBFS data bases, we developed a matrix of potential projects that we think could ultimately be produced by the OBFS network (Figure 3). However, creation of such a matrix requires full and careful inventory of existing OBFS data bases and determination of a unified set of monitoring variables and sampling protocols for systematic long-term monitoring at all or a large subset of OBFS sites. Substantial funding will be needed for this effort.

PROJECT	ECOSYSTEM TYPES							
	1	2	3	4	5	6	7	8
Biodiversity								
Species lists								
Species abundance								
Time trends								
Species/area								
Landscape patterns								
Primary Production								
Annual estimate								
Seasonal patterns								
Time trends								
Landscape patterns								

Figure 3. Example of a matrix of data sets that might be generated through an OBFS data network. Ecosystem types might represent any number of classification schemes encompassing the many different settings of OBFS member sites (all of the biomes in North America are represented by the OBFS network). Cells represent the numbers of field stations that have data at various levels or phases, i.e., most stations would be expected to have species lists; the fewest would have long-term landscape patterns.

Recommended Pilot Projects

We recommend two prototype projects to demonstrate the potential benefits for OBFS sites joining the effort to create a national network for compiling and analyzing environmental data bases and to permit stations to enter the network as individual resources permit.

The first approach concerns the general question of how species richness varies on local to regional scales, based upon species lists currently available at OBFS sites. There are at least four reasons why a compilation of lists of species from field stations would be beneficial. First, the OBFS could determine what proportion of species are represented by field station sites. Second, patterns of species distribution could be illustrated for the United States. Third, the list would represent a benchmark against which further change could be compared. Fourth, the compilation would identify locations where additional data are needed. If a biodiversity data base compiled in this manner is as robust as we suspect, we might be able to fairly quickly begin to relate the influences of various measures of environmental change in ways that are scientifically meaningful, as well as helpful, to biodiversity conservation.

A second project would develop and analyze a basic ecosystem attribute that might be a more robust predictor of environmental change than could be derived from synthesis of species lists. After much discussion, our group decided to focus on primary production in aquatic ecosystems **(we had a table here originally, does anyone remember what it was?).**

A fewer number of stations are expected to have long-term data on aquatic primary production. However, there may be fewer difficulties in comparing data from one site to the next. Again, a simple comparison of annual primary production in aquatic ecosystems from sites throughout the OBFS network could be used to demonstrate the utility of the network. Scaling up, one might compare these long-term records with El Niño events or weather patterns.

Analysis of the Prototype Data Bases

Once species abundance data become available over time, it may be possible to graphically illustrate changing distributions of species abundance response surfaces. Changing distributions could then be contrasted with climatological response surfaces generated by advanced regional global climate models (GCM).

Once patterns of change are documented, the development of testable hypotheses could follow. Some questions developed during the workshop are: 1) does an increase in variance of environmental factors produce an increase in variance in measures of community composition and abundance?; 2) are there changes in the types of responses that are predictable (e.g., decreasing dominance of K-selected species or taxon types with increasing environmental change)? It is possible that data gathered across field stations may be used to test predictions of the consequences of environmental change. Exactly which questions may be examined is beyond the scope of this report.

This effort will be coordinated by Stanford (Flathead Lake Biological Station) and McKee (Andrews Experimental Forest) working with Jim Gosz (LTER Network Office, University of New Mexico).

IV. RESOURCES FOR INFORMATION MANAGEMENT AT OBFS SITES

INTEGRATING SCIENCE AND DATA MANAGEMENT

While field stations have hosted numerous field-based projects for decades, access to these data to answer questions at a regional or national scale is problematic at best. Data sets available at individual stations is not widely known, and the capability to share data across several stations is virtually impossible. Thus, workshop participants discussed technological resources available to facilitate data storage, retrieval, sharing and archiving. Specific examples of data that would have value to all stations if stored in a common format include: (1) a directory of human resources for individual stations, and (2) a catalog of data bases and other types of information available for each station.

Personnel Database

A personnel data base was the first data base implemented by the LTER network. This proved useful to LTER for facilitating networking (i.e., finding others with similar research topics, teaching interests, data sources, etc.). One of the first OBFS network efforts should be a transfer of the successful LTER personnel data base structure to OBFS members. This data base would include name, address, e-mail, areas of interest, areas of expertise, category of research interest, current project titles, status (faculty, administrative, graduate student, visiting scientist) and other information about people and research resources at stations. The existing OBFS directory is a very good start, but a well-developed, web-based, searchable data base is essential to the networking goal.

Basic Information

An OBFS network would include archiving metadata and associated data using the standards described by Michener et al. (1998). Data to be captured by the OBFS network to be made available on the Internet include:

(i) Station Resources (biological, physical), including size, habitat type, latitude, longitude, elevation, freshwater, saltwater access, dive room, boats, classroom facilities, dorms, housing facilities, meal service, specimen collections, etc.;
(ii) Education Opportunity Data base, including list of classes taught, schedule of classes, K-12 programs, internship programs, in-service/continuing education units, volunteer-docent program, seminar series;
(iii) Publication Data base, including papers published from research at stations;
(iv) Species List Data, including lists of mammals, birds, plants, fish, amphibians, reptiles, and insects;
(v) Species Population Biology/Phenology Data, including mammals, birds, plants, fish, amphibians, reptiles and insects; and
(vi) Abiotic/Environmental Data, including meteorology, soil chemistry, water chemistry, hydrology, atmospheric deposition and surface or ground water.

TOOLS AND APPROACHES

Metadata

OBFS stations need tools to handle metadata, which is information about data bases (e.g., who collected it, when, how, etc.). A variety of metadata tools have been developed or are under development at the San Diego Supercomputer Center (SDSC), National Center for Ecological Analysis and Synthesis (NCEAS) and at LTER sites and other locations. Currently, existing tools are site- or topic-specific and do not address comprehensive OBFS needs. Specific examples include the NBII (National Biological Information Infrastructure) MetaMaker program, as well as an Arc-Info macro tool for spatial data. OBFS should participate in providing input to and benefiting from activities undertaken by metadata working groups associated with NCEAS, SDSC and the NBII. Demonstrations, workshops and pilot projects might address these developing issues. Continuing education for OBFS members is a high priority. Opportunities exist for developing a series of workshops that are conducted independently or in association with the annual OBFS meeting for training, metadata content specification and cross-site standardization. [For additional information see Michener et al. 1998, Chapter 8]

Standard Levels of Computational Infrastructure Development at Individual Stations

OBFS data management and networking should reflect the varying needs, infrastructure and capabilities at OBFS stations. There are no standard solutions that will work across-the-board for all sites. Instead, we recognize a continuum of existing capabilities and needs that have to be approached with a wide range of technological solutions.

It is be instructive to establish three hypothetical stages in the evolution of data management and networking at OBFS stations. This approach was initially followed by the LTER Network of sites in developing climate data collection standards and focussed on various levels of implementation that varied by technological capabilities and funding availability. Recently, LTER attention has shifted to developing standard approaches to climate data base management (http://www.lternet.edu/nis), especially data interchange standards. The following discussion is based on a set of data management guidelines developed for the University of Mississippi Field Station planning workshop, a final summary of a data management workshop held in association with the 1997 meeting of the Ecological Society of America (Swain and Michener 1998) and additional input from participants at the 1998 Santa Barbara workshop.

Elements of Effective Data Management

A well-designed and operational data management system can significantly enhance research and administrative productivity at field stations. Successful data management programs typically start at a small-scale and expand incrementally once specific milestones are reached. Several basic tenets that can facilitate successful development of an institutional data management system include:

- start now and start small
- build upon past successes
- keep it simple
- build consensus among users

As with any other component of a research program, data management is highly dependent upon personnel. It may, therefore, be instructive to examine three stages of data management implementation that vary in terms of personnel allocation, capability and technical sophistication (Table 1). Again, most successful data management programs, including those that are most sophisticated, proceed in logical stages from a very simple baseline. The three hypothetical stages used in this example are based on: (1) <1 FTE (full-time equivalent), usually 0.25-0.5 during initial development; (2) approximately 1 FTE; and (3) 2 FTE or more.

Stage 1.

Personnel -- Part-time graduate student, undergraduate student or secretarial support may be employed for Stage 1 tasks. Generally, it is best if one senior staff member administers the data management system (assigns tasks, oversees progress, etc.).

System capabilities -- Data management system responsible for managing institutional data (e.g., lists of site users, site visitors, papers based on research at site, past and on-going research projects, vendors, housing schedules, species lists, etc.).

Technical description (hardware/software) -- A basic data management system for handling institutional data can reside on a standalone PC. Software may initially include a word processing package (e.g., MS WORD) and a spreadsheet (e.g., EXCEL). As data bases grow, PC-based data base management software (e.g., PARADOX) can be added. Management of reference lists may be accomplished using one of several bibliographic data bases (e.g., REF11, Papyrus, etc.). Initial computational environments typically do not have a high degree of network connectivity, as network maintenance requires higher skill levels and personnel commitments.

Administrative issues -- It is generally desirable to prioritize data management tasks based on their potential utility to the greatest number of users. Often, sites benefit from having a steering committee (3-5 individuals, including one or more scientists and administrators, i.e., the site stakeholders) that helps establish policies, priorities, and procedures. Specific administrative issues that must often be dealt with during Stage 1 include: (1) prioritizing data management tasks; (2) developing site and data acknowledgment policies; (3) developing policies for obtaining reprints based on work done at site; and (4) initiating a long-term development plan.

Stage 2.

Personnel -- Assigning a full-time individual to data management can greatly facilitate system development and enhance functionality. A recently graduated biology student with computational skills/interests may be sufficient for a developing data management program.

System capabilities -- As the computational environment (hardware and software) expands at an institution, there is an increasing need for intra-site networking (e.g., Local Area Network), hardware and software trouble-shooting and more sophisticated software (e.g., a DBMS, such as Paradox or Access, SAS or SYSTAT for data analysis and graphics, etc.). A modest LAN can facilitate data sharing and archival, as well as sharing of peripheral devices (printers, plotters, etc.). Often, personnel associated with the data management system are involved in establishing and maintaining a climate station (and resulting data base) and connecting laboratory instrumentation to PCs. During Stage 2, attention is often devoted to developing institutional data bases (e.g., climate, water quality, etc.) that may be useful to the broad community of site users (university and visiting scientists, graduate students, etc.).

Technical description (hardware/software) -- A site in the second stage of development frequently has several PCs, printers, one or more plotters and archival hardware (read-write optical disk, tape, etc.). Offsite communication may be accomplished using high-speed modems and dial-up local access to an Internet provider.

Administrative issues -- Formal archival (back-up) procedures should be adopted for institutional data bases. Attention should be paid to developing data catalogs (reference list of site-specific data sets) and metadata (data documentation) guidelines and protocols for high priority data sets. Several critical issues generally arise during Stage 2 that deserve concerted attention and discussion including: (1) desirability of a site GIS (geographic information system); (2) need for a dedicated communication line (e.g., T1) for constant Internet access; (3) need for fully-functional data base management systems that are based on structured query languages DBMS (e.g., ORACLE, etc.); (4) need for multiple hardware platforms (PCs plus UNIX stations); and (5) need for software, especially statistical, support. Personnel requirements associated with these system enhancements should be thoroughly addressed. For example, GIS implementation can be especially costly for a developing station and may detract resources and attention from more critical tasks. Salary and personnel issues generally arise during Stage 2. As data management systems become more sophisticated, there is an increasing need for more highly skilled employees. Frequently, data management personnel turnover at field stations becomes an issue requiring administrative attention.