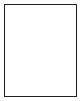
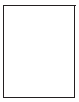
Proceedings of the 1993 LTER Data Management Workshop

Editors: Rick Ingersoll, Rudolf W. Nottrott and John H. Porter

Participants and Contributors: K. Baker, B. Benson, C. Bledsoe, J. Brunt, J. Briggs, G. Calabria, J.Hastings, D. Henshaw, R. Ingersoll, J. Jefferson, D. Jones, T. Kirchner, M. Klingensmith, L. Krievs, R. Lent, M. MacKenzie, T. Maddux, B. Meeson, E. Melendez, B. Nolen, R. Nottrott, J. Porter, S. Stafford, J. Vande Castle, C. Veen



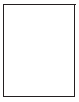
Long-Term Ecological Research Network (LTER)



Distrubuted by the Long-Term Ecological Research Network Office
College of Forest Resources, AR-10
University of Washington, Seattle, WA 98195
Phone: 206-543-4853 or FAX: 206-543-7295

For instructions on access to electronic information on LTERnet (including this report) send e-mail to
Overview@LTERent.edu or
Helper@LTERnet.edu

Supported by a grant from the National Science Foundation, Division of Environmental Biology (93-00679). Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.



Long-Term Ecological Research Network
College of Forest Resources, AR-10
University of Washington
Seattle, Washington 98195

For instructions on access to electronic information on LTERnet (including this report) send e-mail to
Overview@LTERent.edu or
Helper@LTERnet.edu

Proceedings of the 1993 LTER Data Management Workshop

Table of Contents

Proceedings of the 1993 LTER Data Management Workshop

Editors: Rick Ingersoll, Rudolf W. Nottrott and John H. Porter
Participants and Contributors: K. Baker, B. Benson, C. Bledsoe, J. Brunt, J. Briggs, G. Calabria, J.Hastings, D. Henshaw, R. Ingersoll, J. Jefferson, D. Jones, T. Kirchner, M. Klingensmith, L. Krievs, R. Lent, M. MacKenzie, T. Maddux, B. Meeson, E. Melendez, B. Nolen, R. Nottrott, J. Porter, S. Stafford, J. Vande Castle, C. Veen

1.0 Executive Summary

The 1993 Long-Term Ecological Research Data Management Workshop was held July 29-31 in Madison, Wisconsin. A total of 25 data managers and other interested scientists attended the workshop, including at least one representative from each LTER site (except for CDR). The workshop focused on two major areas. First, reports were presented on activities coming out of the 1992 workshop. Secondly, working groups were established to formulate plans for the coming year and to address issues of LTER-wide concern. A guest speaker from NASA's Pilot Land-Data System provided valuable insights into methods of data management used in another organization.

1.1 Data Management Activities since the 1992 Workshop

1. The **symposium "Environmental Information Management and Analysis: Ecosystem to Biosphere Scales"**, organized by three LTER data managers, was held on May 20-22, 1993, in Albuquerque.
2. A prototype **LTER all-site bibliographic database has been developed.** The bibliography uses data from 15 LTER sites. The prototype will serve to demonstrate the benefits of the database and to resolve technical details relating to access and maintenance methods.
3. **Relationships with the Chinese Ecological Research Network were strengthened**, with three LTER data managers participating in an extended visit to China.
4. Following widespread discussion of interactive data access (IDA) systems at the 1991 and 1992 workshops, **an increasing number of sites have implemented some form of on-line data access system**.

1.2 1993 Workshop Accomplishments

1. A **draft plan for expansion of the LTER Catalog of Core Data Sets** was completed. The expanded catalog will include all long-term datasets at the sites, and calls for the **recovery of historic datasets.**

2. A **set of minimum standards for metadata content was adopted**. These standards will be implemented in a common data exchange format for metadata, which continues to be under development.
3. The workshop participants agreed on a **plan for a data managers meeting in 1993 with an extended agenda**, including participants from other organizations with data management problems similar to or related to those of the LTER network. After one day of discussions on subjects of interest to LTER only, the focus of the meeting will shift to areas where LTER data managers can cooperate with their colleagues in other organizations, and an effort will be made to identify specific projects for collaboration. The meeting will be hosted by the Network Office in Seattle in September 1994.
4. **A "think-big" exercise was conducted** to identify areas where major advances in ecological information management could be made. Areas identified included **improved connectivity** and the **development of improved interfaces** with both human and technical components.
5. Recommendations were formulated to assist in preparation for site reviews.
6. Gopher and WAIS **network information servers were demonstrated** and their potential roles in LTER information management were discussed.

2.0 Activity Reports

2.1 Chinese Ecological Research Network, CERN (James Brunt, SEV)

Five data management professionals from the Chinese Academy of Sciences - Chinese Ecosystem Research Network will be in the U.S. later in 1993 to participate in a collaborative research data management training course at the University of New Mexico. The trainees, Liu Degang, Jiang Hong, Su Wengui, Liu Guobin, and Liao Sunbao were selected for participation during a 1-day data management workshop in Beijing in September of 1992. Upon returning to China, these individuals will become responsible for major aspects of CERN data management including training. A number of LTER data managers will be participating in the 3-month course. U.S. funding for this training was provided by the U.S.-China program, International Programs at the National Science Foundation.

2.2 Environmental Information Management Symposium (James Brunt, SEV)

The symposium entitled "Environmental Information Management and Analysis: Ecosystem to Global Scales" was a success. Held at the University of New Mexico in Albuquerque in May, the symposium attracted over 150 outside participants to hear and discuss over 35 papers. The symposium began as a working group discussion at the LTER data management meetings in Toronto in 1989. A proposal submitted by Bill Michener (Joseph W. Jones Ecological Research Center), James W. Brunt (SEV), and Susan G. Stafford (AND) was funded by the National Science Foundation in 1992. Call for papers went out in August of 1992. Final edits for an edited volume made up of the accepted papers are in progress. The volume will provide in-depth information on topics heard in the symposium

as well as some new information. LTER participants have quite a few papers to be included in this volume which is on the fast track with Taylor and Francis, New York for publication early in 1994. Taylor and Francis has a good history of publication in the environmental arena.

As a trial, unedited first drafts of most of the symposium papers were put up on a WAIServer immediately following the symposium for access via Internet. People interested in reviewing any of the documents were encouraged to access the papers and provide critical reviews to the editors.

Reactions to the format and content of the symposium were very positive and the most frequently asked question was "When is the next one?". Organizers and participants discussed the possibility of another meeting in 2-3 years.

2.3 NASA Collaboration (John Vande Castle, NET)

John Vande Castle reported on the LTER/NASA-EOS interactions as it will relate to site and Network data management. This collaboration has developed from past discussions between LTER and NASA collaborators and will lead to a number of new initiatives. A summary of current initiatives is described in reports from meetings at Sevilleta and Goddard Space Flight Center which were distributed to the sites and are available on-line at LTERnet.edu via FTP or Gopher. An overview of future collaborations will be discussed at a LTER/NASA workshop at the upcoming LTER All Scientists Meeting.

The newer initiatives involve collaboration with the earth science groups at NASA, including those involved with specialized or newer sensors such as AVIRIS, and MODIS, global change programs involving Landsat-TM and AVHRR data, and aircraft remote sensing projects.

There have been initial collaborations to acquire Landsat-TM data for all LTER sites, with 12 acquisitions by NASA in 1992 and 26 planned for 1993. Access to these and other datasets is complicated by the magnitude of these kinds of data and licensing constraints (e.g., EOSAT does not allow shared distribution of Landsat-TM data). As the collaboration proceeds, the role of the Network Office, the LTER sites, and NASA (including the NASA/EOS archive centers) will have to be defined with respect to data access and management.

In discussion it was suggested that good communication was essential as this collaboration evolves. John suggested that a report of the current status of remote sensing activities at each site was important. Although some information for each site is contained in the Sevilleta meeting report, a current status is needed. The compilation of this report will be arranged with the data management group after the September LTER All Scientists Meeting.

2.4 Data Management Policies (Karen Baker, PAL)

A data management policy is important for LTER sites because it defines both the responsibilities and rights of the investigators, data providers, and data users. The LTER sites are in various stages of development of individual site data management policies. The needs of each site will vary so data management policies may be expected to range from an

informal unwritten policy to multi-page formal documents.

A data management policy can include a description of the data administration (e.g., personnel, costs, policy interpretation), data types (e.g., published and unpublished data, metadata, and individual investigator data), data restrictions, data availability, data storage and access procedures, investigator responsibilities (e.g., data entry, QA/QC), and user responsibilities (e.g., acknowledgments).

There are several valuable references for developing a data management policy. Organizations or projects responsible for gathering large amounts of data have project-wide policies (e.g., WOCE/telnet delocn.udel.edu/USER=INFO) or deal with datasets individually (e.g., GLIS/telnet glis.cr.usgs.gov). The LTER ad hoc committee on data access, in conjunction with the LTER Network Data Manager, has developed a general set of guidelines which is available on-line through the LTERnet Gopher system. A survey paper by John Porter and Tom Callahan, presented at the international symposium (Environmental Information Management and Analysis: Ecosystem to Global Scales) held in Albuquerque, NM, in May 1993, provides a summary of existing LTER site data management policies. Each LTER site has been asked to provide an electronic copy of its site data management policies to the Network Office, so that it is accessible to all of the other sites. Given the dynamic nature of these policies, this archive will be useful both in the initiation and revision of policies.

2.5 Intersite Climate Database Project (Tom Kirchner, CPR)

The Intersite Climate Database project was put together by John Gorentz at Kellogg Biological Station to develop prototypes for methods to exchange data between LTER and other sites. John's approach centered on using e-mail as the transport mechanism for exchanging data. The data are accessed from SQL databases using procedures that accept requests sent to them from e-mail, interpret the requests, obtain the data from the server, package the data as a mail message and returns it. John Briggs (KNZ), Barbara Benson (NTL), and Mark Klingensmith (BNZ) participated in the project by assembling their sites climatic data into an SQL database.

A second thrust of the project was to put together a library of interprocess communication (IPC) functions to implement file transfer procedures directly using remote procedure call (RPC) methods. These functions were designed to facilitate the construction of distributed client and server applications. The applications could be written in either FORTRAN or C. Tom Kirchner (CPR) took responsibility for this part of the project.

John Gorentz was not present at the meeting and thus could not present a status report about his progress on the creation of the e-mail based system. However, the participants who were responsible for making available climate databases confirmed that the databases have been completed and are now accessible.

Tom Kirchner reported that the IPC library has been completed and is now in the last stages of being documented. An application called ltermenu has been developed to allow interactive access to datasets. Ltermenu enables the user to select a dataset, display the metadata for the dataset, plot or list data in the dataset, and download data and metadata files. The ltermenu program is based upon the CPR style of metadata files. Ltermenu has been ported to X-Windows under Unix and to the PC under DOS using the PC-NFS Tool-

kit from Sun, although some bugs are still being fixed. A server designed to access the KNZ climate database is now being developed. The server will accept a request for a dataset (table) from ltermenu or similar applications, retrieve the data, generate a metadata file to describe the format of the data, then send the metadata file and data to the client application.

2.6 Publication Forums (Eda Melendez, LUQ)

One of our main objectives as data managers is to provide new research insights by generating new graphical relations for looking at data and better ways of combining data. Routinely we get involved in activities to look for better ways to perform these tasks at an individual level at each site, and at the network level as a group. Sometimes we find solutions to our individual problems, and we share them among ourselves at our workshops. Our work of managing, analyzing and graphing data sometimes bring out specific characteristics of the data that give a better understanding of the subject to the investigator. In the process we discover new methods to perform these tasks that are distinctive for the kind of data we manage.

The work has been done at each site. It is clear to us that we need now to communicate our findings. Furthermore, in addition to our group, others are starting to envision Data Management as a scientific discipline in its own right. In 1992 at the Hawaii workshop, James T. Callahan of the National Science Foundation, pointed out this fact and mentioned that information management must deliver what is needed to other scientists and that we should aim at objectives.

With this motivation in mind, we have started to prepare a list of journals that can be our media to show our link to scientific discovery. Finding out which of the material we have in our hands is publishable and establishes that link is the first and for some the most difficult task to achieve. The following is an initial list of journals and bulletins:

1. Ecological Applications
2. Envirometrics
3. Academic Computing
4. Environmental Science and Technology
5. IEEE Computers
6. Antarctic Journal (regional)
7. Sigmod Bulletin
8. ESA Bulletin

2.7 Tour of the Environmental Remote Sensing Center at (ERSC) the University of Wisconsin-Madison (Mark MacKenzie, NTL)

LTER data managers participating in this summer's meeting were given a tour of the facilities at the Environmental Remote Sensing Center (ERSC) at the University of Wisconsin-Madison. The mission of ERSC is to provide a physical and intellectual focus for interdisciplinary research on the development and application of remote sensing in natural

resource management and environmental monitoring. By design, ERSC's research agenda is broadly based: 1) disciplines range form archeology to zoology, 2) problems range from basic to applied and from technique-oriented to application-oriented, 3) research sites span the local to global continuum, and 4) research efforts typically involve the merger and synthesis of remotely sensed data with other sources of land-related information in a GIS context. ERSC is administered by the UW-Madison Institute for Environmental Studies (IES). ERSC is the physical focus for advanced remote sensing and GIS courses taught in support of IES's Environmental Monitoring Program and also serves numerous other academic programs. ERSC is also an active participant in the Spatial Information and Analysis Consortium (SIAC) at UW-Madison. SIAC members are drawn from various colleges within UW-Madison. The mission of SIAC is to provide campus-wide leadership and coordination in advancing the theories and methods of collecting, managing, analyzing, and applying spatially referenced information about the natural and cultural environment.

Participants in the tour were shown the facilities available at ERSC. These facilities include research computational resources including Unix workstations and DOS based PCs supporting ERDAS, Arc/Info, and McIDAS. Instructional computational resources include DOS based PCs. All computational resources within ERSC are connected via a local area network (LAN) which is connected via ethernet to the campus backbone and from there to Internet and other wide area network (WAN) links. Through these WAN links, ERSC has been able to utilize peripherals and resources available at other state and federal agencies (i.e., Wisconsin Department of Natural Resources, USGS, LTER Network Office). Ancillary peripherals at ERSC include a microdensitometer, digitizing tablets, film recorder, optical and magneto-optical disk drives, and various printers and plotters.

The tour also included a demonstration of ERSC activities directly related to NTL-LTER research efforts. This demonstration illustrated applications of remote sensing and GIS to ecological research including land cover classification, remote sensing of limnological parameters (i.e., temperature and turbidity), and digital elevation models.We are planning to make parts of this demonstration available for viewing through the Network Office Information System.

I would like to acknowledge Frank Scarpace, Pete Weiler, Marcia Verhage, Larry Seidl, and Randy Wynne for their contributions to the ERSC tour.

3.0 General Topics

3.1 Network Information Servers (John Porter, VCR and Rudolf Nottrott, NET)

Gopher and WAIServers are increasingly in use at LTER sites. Gopher information servers can provide public access to textual and graphical materials via easy-to-use menus. WAIServers provide rapid searching capabilities, both alone and inside Gopher menus. The advantages of interactive data access systems were enumerated in the 1991 workshop proceedings and a survey of existing LTER interactive data access systems was presented in the 1992 proceedings. The ease of operation of Gopher servers has greatly expanded the list of LTER sites with interactive data access systems in the past year, with 8 sites operat-

ing Gopher or WAIServers.

A demonstration of the LTERnet Gopher system was presented at the workshop. The demonstration included displaying text from LTER Network Office publications, searching the on-line Personnel Directory, the Core Data Set Catalog and linking to other LTER Gopher systems.

The wide use of such systems to provide access to information has many advantages, but also poses challenges. Providing mechanisms to update information in the on-line system so that it remains current is one of the most critical challenges. This challenge is closely related to the following questions:

1. What information should be provided?
Gopher information servers provide excellent access, but have few options for security. This makes them suitable for information about sites, site publications and site activities, but not necessarily for data itself, as there are often proprietary constraints on the distribution of recently collected data. Similarly, the servers provide data on what information was viewed from a given host computer, but not the identity of the individual doing the viewing. Thus, it may provide inadequate information to track database use for some classes of data.

2. Where should information reside?
This question relates to the ease and mechanisms of update. For example, if each site operated its own Gopher server and maintained a personnel directory for the site, the LTER Network Office could create a master directory simply by providing links to the individual site directories. Alternatively, the LTER Network Office could maintain a master personnel directory of all sites and then individual sites could link to it from their Gopher servers. Both approaches have advantages and disadvantages. The first approach requires that every site maintain a Gopher server and store personnel information in a standard form. The second approach requires that the Network Office database be updated -- often by people who are unfamiliar with the personnel at the site.

These questions were raised at the meeting, but further experience with the different modes of operation will be needed to establish criteria for making decisions about how to handle different kinds of data. It was noted that Gopher-based informations systems are not information management systems per se. Their strengths are greater for browsing descriptions of data more than for accessing or manipulating data. It was agreed that presenting a demonstration of existing interactive data access systems would be an important addition to the upcoming LTER All-Scientists Meeting.

3.2 Information Management in the NASA Pilot Land-Data System (Presentation by Blanche Meeson, NASA)

Blanche Meeson, Project Manager of the NASA Pilot Land-Data System, addressed the workshop. Blanche noted that NASA deals with many types of data, from satellite to laboratory measurements, from point to 3D, from global to local with times varying from subseconds to yearly, or seasonally, and that some data have no temporal or spatial tags. The volumes of data number in the terabytes, but the “granule” sizes vary from gigabytes from

satellites to kilobytes from aircraft, ground, and laboratory measurements.

NASA information systems vary in their structure depending on the purpose of the systems and the types of data they handle. Processing and analysis systems focus on the analysis of a limited number of datasets. Science Project systems focus on integration of diverse datasets, with a broad spectrum of data sources and spatial types, and a large number of datasets (30-100). Distribution/repository systems are divided into Flight/Science Project databases where there are a limited number of users and datasets and metadata is limited because only experts on the instrument are funded and have access to data. On science projects a limited number of users and large number of datasets are expected. For discipline-level databases there is a potentially large number of users and datasets incorporating a high level of diversity. Composite systems attempt to play many roles. An example of such a database would be EOS-DIS, which is trying to merge all these types of databases.

Blanche discussed how the complexity of management increases with number of datasets and the diversity of the user community, the role of standards and guidelines, and the need to balance top-down and bottom-up perspectives. She also discussed the role of metadata in producing datasets that can meet the National Academy of Sciences "20-year test", which maintains that knowledgeable scientists should be able to make sense of the data 20 years into the future and the need for flexibility in developing standards. She provided detailed examples of dataset abstracts, and metadata. The critical aspect of such documentation is that it be sufficient to let a scientist believe that the data is accurate enough for his/her use. She also emphasized that creation of metadata needs to be prepared up front. Post hoc reconstructions have not worked well in the past.

Blanche also discussed the issues unique to spatial data and some of the ongoing attempts at NASA to arrive at general data formats. She emphasized that final, or even good solutions that integrate data and metadata, with the capability to analyze data have not been reached.

3.3 All-Site Bibliographic Database (Rudolf Nottrott, NET)

The completion of the initial phase of the LTER All-Site Bibliography Project, and the implementation of a prototype on the LTERnet Information Server, mark major steps toward the goal of providing Internet users with access to bibliographic information from all sites. Development of the prototype, which is located at the LTER Network Office in Seattle, provides partial achievement of this goal. The ultimate goal could be the development of a distributed database, with each site maintaining a local version in a standard format, searchable on-line from any of the other sites.

For the prototype, Harvey Chinn, under the guidance of LTER Research Coordinator Caroline Bledsoe and with the assistance of the site data managers and Rudolf Nottrott at the Network Office, compiled text versions of the bibliographic databases for 15 sites and wrote a set of programs to rearrange them into a common format.

The prototype bibliography at LTERnet is WAIS-indexed and is accessed through a Gopher menu. Searching the bibliography is analogous to a search of the Core Data Set

Catalog or the Personnel Directory. Access to and searching of the bibliography were included in the demonstration of the LTERnet Gopher system at the workshop.

Potential uses of the all-site bibliographic database for individual sites, the LTER network as a whole, and the wider ecological research community are wide-ranging. Multi-site investigations and synthesis of research efforts by LTER scientists are facilitated by improved accessibility to bibliographic information. With research information more accessible, LTER scientists can more easily plan multi-site experiments and synthesize information. Scientists from the wider ecological research community can compare their data or sites to similar LTER data or sites, developing more robust, general ecological principles. A bibliographic database can be cross-linked to the LTER Core Data Set Catalog, providing researchers with access to citations, as well as descriptions of specific datasets. In addition, agency administrators, decision makers and managers will have a very large body of ecological information identified and indexed for their use.

3.4 Site Review Tips (Susan Stafford, AND)

Several of the LTER sites either have had, or soon will be having, a NSF LTER site review. There was much interest in sharing “tips” that would make it easier for the newer sites to prepare for their upcoming data management reviews as well as sharing experiences among the veteran sites. The following pointers were discussed:

1. Be factual.

2. Be honest.

3. Be candid - if you are comfortable with the review committee and your fellow PIs. Otherwise, simply tell it how it is and don't get defensive.

4. Present the “Big Picture”.

5. Present the philosophy of data management at your site.

6. List the resource allocation priorities. Include budget information if available.

7. Consider a limited number (perhaps 3) of key points that the reviewers should take home with them. Time will not permit an in-depth discussion of everything so you will have to focus on the main items at your site.

8. Time will be limited. Therefore, avoid inadequately tested presentations or those that require more than the allotted time.

9. Although you might be the only individual that “sticks” to his or her allotted time, this will be a demonstration of professionalism and efficiency.

10. Provide, in a hardcopy form, a concise summary of your site’s data management philosophy, e.g., a 2-page “handout” that lists the motivations for data management, the objectives, and the strategies for achievement of those objectives. Include a flowchart and keep it simple. Note that the review team will have at most 1-2 people who are experienced in data management. It should NOT read like a computer manual. This material should be included in the packet of advance material sent out to the review team several weeks ahead of the visit.

In terms of total time, expect 15 to 45 minutes. Be sure to leave time for questions. Inter-

ruptions will happen! One model is 15 minutes of presentation, 15 for questions, and 15 for a “walk-around” tour/demonstration. Be flexible and try to adapt your presentation to the interests of the individuals comprising your review team. One-on-one demonstrations can be arranged as necessary on an ad hoc basis. One site invited the reviewing data manager to come in one day earlier so that there would be adequate time for touring facilities and discussion.

Know your review team. Some may need to be educated about the LTER program and data management in particular. Also, be sure to invite local folk who might have an interest in acquiring more knowledge about data management. The review of your data management program can provide an opportunity to educate the “greater community”.

One of your objectives should be to demonstrate that data management is an integrated component of the research being conducted at your site. Other investigators can assist by including the data management aspects of their research into their presentations. Other PIs will continue to be our best ambassadors for a strong site data management system. We want to portray the status and stature of data management as positive.

Don't be surprised if you know the person reviewing data management. Other LTER data managers are routinely asked to serve in the review capacity. Some have provided a set of questions to be answered ahead of time. This is very helpful.

Keep in mind that these reviews are designed to be helpful but because of time constraints, the feedback session will be short. There will be a number of issues that will be raised. Review teams feel obligated to find more negatives than positives - it's human nature. Don't take it personally.

3.5 Data Management Infrastructure/NSF (General Discussion and Working Group)

Committee: Eda Melendez (LUQ), Cindy Veen (HBR), Jordan Hastings (MCM), Karen Baker (PAL), Richard Lent(HFR), Jack Witham (Holt Research Forest - guest), and Susan Stafford (AND)

Dr. Jim Gosz, Division Director of Environmental Biology at the National Science Foundation (and on leave from the SEV site), spoke to the group about opportunities for future information infrastructure to support research and science. The LTER continues to be a flagship for NSF and we should seize the opportunities in this arena as well. As a spin-off of this discussion, a working group convened for the express purpose of prioritizing needs that could be satisfied with supplemental funding for data management within LTER. Jim Gosz had challenged the group to be creative in determining such needs at the network level, as opposed to equal distribution of funds among 18 sites.

The working group agreed that the most effective use of such resources would be to concentrate the investment on a set of critical capabilities so as not to dilute the dollars. Thus, if the goal was to achieve some common level of capability and functionality across the network, then some of the more advanced sites might not benefit directly from these resources. Nevertheless, there would be indirect benefits, because their colleagues in the scientific community would be better connected and thus their expertise would be avail-

able to the whole network.

NSF is in the business of supporting science --- good science. Consequently, development of infrastructure must be in the context of higher quality research and higher quality research products.

There are three predominant themes in ecology today:

1. Issues of scale
2. Spatial issues
3. Issues addressing social dimensions (co-directed research)

We recommend that requests for infrastructure improvements be connected to at least one of these major themes.

With regard to the first theme, we identified two high-priority items.

Complete connectivity to the Internet for all sites, including field stations, was given highest priority. From Jim Gosz's comments it was not clear whether or not this was a given. In any case, this is the first order of business. Sharing data across the network needs to be as routine as reading our email and this will require full connectivity among all sites and field stations.

The second priority was for improvement of interfaces -- hardware and software, as well as “human”. Workshops to develop and share technology and solutions to common problems were considered to be of crucial importance.

Third, the following group of capabilities were also identified:

1. Distributed modelling
2. Restoration of older datasets
3. Spatial metadata

One idea for testing the reliability of older datasets was to develop a regional model that would test the congruence of older datasets with the simulation results. There was much discussion on the value of historical datasets as well as the great difficulty in acquiring the resources to complete the all-important cleaning and housekeeping details on the associated metadata.

Furthering collaborations with industry was identified as a means to help build LTER data management infrastructure. Mechanisms need to be identified whereby solutions to ecological information problems have marketing appeal to hardware/software developers. Currently, CEOs haven't been convinced that the profit margin was large enough to warrant their time and energy in forging collaborative efforts with the ecological scientific community. We see this as changing in the near future as talk in Washington progresses on National Information Infrastructure (NII) and Information Infrastructure Technology Assessment (IITA).

4.0 Working Groups

4.1 Infrastructure, Future Meeting Formats, and Outreach Activities

Committee: Rick Ingersoll (NWT), Barbara Benson (NTL), John Briggs (KNZ), James Brunt (SEV)

Redefinition of the mission of the Data Management Task Force (DMTF) and **formation of a Data Managers' Meeting Organization Committee** (DMMOC) were suggested in order that (1) routine, network-wide tasks be equitably distributed amongst the LTER information managers and (2) all information managers have the opportunity to become involved at the network level.

The primary functions of the DMTF are to ensure (1) continued funding for future Data Managers' Meetings, which might require proposal generation, and (2) that an information manager attends the LTER Coordinating Committee meetings. The DMTF should not be responsible for the organization of future Data Managers' meetings. The DMTF currently consists of Barbara Benson, John Briggs, James Brunt, Tom Kirchner (CPR), John Porter (VCR), Susan Stafford (AND), and Rudolf Nottrott (NET, ex-officio). The 6-member format (with at least 2 members having signatory powers; plus the network information manager as an ex-officio member) was deemed satisfactory but a mechanism was proposed that would provide all information managers with the opportunity to serve on the DMTF. Specifically, any information manager may volunteer for the DMTF and, if more than 2 individuals do so at one time, an election can be held to select the 2 candidates, who should expect to serve at least 3 years.

The primary functions of the DMMOC are (1) to organize the Data Managers' Meetings, and (2) to edit and produce reports from those meetings in a timely fashion. This committee will consist of 6 volunteers: 3 members of the DMTF plus 3 "at large" members; care should be taken that it is not gender-biased. Volunteers for this committee were Barbara Benson, Rick Ingersoll, David Jones (ARC), Mark Klingensmith (BNZ), Rudolf Nottrott and Susan Stafford. The chairperson will select co-chairs so that the tasks necessary for the smooth preparation, operation, and "wrap-up" (i.e., report production) of the meeting are equitably distributed.

It was recommended that the information managers, as a group, make more effective use of electronic mail and LTERnet during the course of the year so that our meetings are more efficient and productive. The formation of a "data managers only" group on LTERnet and more extensive use of surveys would facilitate this. Specifically, much of the discussion for working group topics chosen for the subsequent meeting should be carried out over email so that the reports for those topics can be presented at the meeting. This would allow more time for presentations and invited speakers, as well as expediting final report production. There would still be the option for introduction of additional topics during the "site flashes."

The 1994 Data Managers' Meeting is planned for Seattle in the autumn of that year. We will evaluate further the advantages and disadvantages of (1) meeting at LTER sites

(rather than in conjunction with ESA), (2) non-summer timing of meetings, and (3) an expanded 3-day format with increased participation from non-LTER information managers or those in related fields. Potential advantages include (1) opportunity to visit other research sites and data management laboratories, (2) scheduling of the meeting for less hectic times of the year, and (3) increased interaction with non-LTER groups. The primary disadvantage is the increased expense of a meeting separated from ESA. The added expense is not limited to funds and must include time considerations as well.

We need to initiate (and soon) solicitation of interest in the 1994 meeting from groups that include, but are not limited to, the Organization of Biological Field Stations (OBFS), the Land Margin Ecosystems Research program (LMER), the Southern Association of Marine Laboratories (SAML), the Long-Term Research in Environmental Biology program (LTREB), NASA, Oak Ridge National Laboratory (ORNL), the J.W. Jones Ecological Research Center, and the US Geological Survey (USGS). The individual sessions should concentrate on areas where LTER data managers can cooperate with their colleagues from the other groups (e.g., metadata, on-line information services, QA/QC, scaling) and both LTER information managers and invited speakers could make presentations. An effort should be made to identify specific projects for collaboration. At least one day (preferably the first day), however, should be devoted to LTER business.

4.2 Development of Long-Term Data Set Catalog and Restoration of Historic Data Sets

Committee: Rudolf Nottrott (NET), Mark Klingensmith (BNZ), Barbara Nolen (JRN), Blanche Meeson (NASA - guest)

The LTER Catalog of Core Data Sets has been a valuable tool in providing metadata on core datasets -- the datasets relevant to research in the core areas defined by the LTER network and NSF. However, this catalog with its 250 entries represents only a fraction, perhaps as few as 10 percent, of the datasets maintained at the LTER sites. **Expansion of the catalog to include all long-term datasets at all sites is a desirable and realistic goal**.

When our discussion covered the area of historic datasets that may require considerable effort to bring them into an “exchangeable” state, participants realized that the scope of any catalog expansion project should be broadened to include restoration of high-priority historical long-term datasets. After their restoration, these datasets would be included in the expanded catalog.

The Core Data Set Catalog, originally produced in hardcopy form, lists 250 datasets from 18 sites. The on-line version of the catalog, located at LTERnet, is searchable directly over the Internet, or by electronic mail. It will form the starting point for our efforts to create a catalog of all long-term datasets.

The data managers developed the following **draft plan for catalog expansion**:

1. At each site, **compile a list of site-related datasets that have been collected or are intended to be collected for more than 5 years**. Include historical datasets (pre-LTER) in the list. For each dataset, list the following information:

Dataset title

Name of investigator(s)

Is the dataset administered by LTER personnel (or with LTER funding) or not (Yes/No)

Is the dataset ready for exchange (Yes/No), i.e., quality assurance/quality control are complete, documentation is complete, dataset is in digital format -- same criteria as datasets in present Core Data Set Catalog.

The data managers will have a **draft list for their sites by the LTER All Scientists Meeting** or shortly after.

2. If a dataset is NOT ready for exchange, then **estimate the cost** for its restoration in terms of:

   Amount of money

   Amount of time

   Number of person-hours

3. For every site, **prioritize** the restorable long-term datasets on the basis of:

   Value of the dataset to the site and to others, in particular its value for intersite research in the LTER network

   Estimated cost of restoration

4. **Determine potential sources of funding** for compiling the complete catalog information for all long-term datasets administered by LTER as listed in 1., and to cover the estimated restoration costs for historic datasets with high priority.

   1-3 should be completed by early 1994 (January/February) so that we are ready to look for funding sources by that time.

4.3 Data Management Curricula Resources (David Jones, ARC)

Development of curricula for use in teaching the application of information science to long-term ecological datasets was discussed, but tabled as a working group. Experience with the Chinese Ecological Research Network and classroom activities from some of the data managers may form a good basis from which to work. Blanche Messon (NASA) commented that it was difficult to teach concepts in this context. Interest in pulling together a series of student-related projects from several sites was expressed, and may become a working group in the future. The GIS examples in the Idrisi project may be a good format on which to base this.

4.4 Data Exchange

Committee: Tom Kirchner (CDR), Don Henshaw (AND), Jeff Jefferson (NIN), Lolita Krievs (KBS), Gil Calabria (CWT), John Porter (VCR), Troy Maddux (SEV)

The data exchange working group continued a process initiated at the 1992 workshop. Its focus is on the development of a common exchange format for metadata (data about data)

so that data can be easily shared among sites. Each site would continue to manage its documentation with site-specific procedures, but would develop “filters” which translate their metadata into the common exchange format.

During the past year, Tom Kirchner has continued development of a flexible format for metadata and a list of minimum documentation standards were developed and were further refined at the workshop. Development of such standards is a necessary precursor to implementation of a standard format because programs using the format need to have some expectation as to the types of information that will be available.

Discussions of the working group centered on the need to identify good test cases for use of the format. There was also considerable discussion about the relative merits of specific formatting options, specifically Kirchner’s “attribute/value” format and standard generalized markup language (SGML). Discussions centered on the need for any format to be machine independent, human readable, and for the format to be able to accommodate hierarchical data structures so that referential integrity would not be compromised by multiple inputs of the same data. It was noted that software systems with such capabilities are easily interconvertible, so that choice of a specific system is not as critical as it would be otherwise.

The group identified **four main objectives**. The first was to develop examples of data exchange and some basic utilities. The second was to review the list of proposed standard documentation items for the minimum standards. This objective was accomplished during the working group section. Thirdly, the group should identify LTER researchers interested in participating in a trial implementation of the system. These activities would include parser development, creation of sample datasets, development of a “template” for the data exchange format, and development of interactive data display and extraction tools based on the format. Finally, it was decided that a proposal to NSF should be submitted by interested sites to pursue full development of the data exchange format. Critical to the success of such a proposal will be a link to a research project that would greatly benefit from such a format.

5.0 Appendices

5.1 Site Flashes

5.1.1 AND (Don Henshaw)

The Andrews had their NSF review in June. Tom Callahan (NSF) led a team of eight reviewers that included Bill Michener. The Quantitative Sciences Group (QSG) and the Forest Science Data Bank (FSDB) was commended in the area of data management in the opening remarks of the review team. QSG and FSDB materials sent to the review team are being packaged into a brochure so that available, ready-made packets will be on hand for future reviews and other requests for information.

The QSG has also reached an agreement with the Bureau of Land Management (BLM) projects located within the Forestry Sciences Lab. BLM has agreed to pick up funding for our assistant LAN manager, as well as a half-time spatial statistics position and a half-time GIS programmer. Pooling resources across projects (LTER, OSU, USFS, BLM, COPE) enables individual projects to pay their fair share of the total costs while having access to the whole collection of services and expertise. This approach helps relieve pres-

sure on a single project such as LTER, as it would be cost-prohibitive for individual projects to underwrite the complete spectrum of statistical and computer services available through the QSG.

LTER investigators Fred Swanson, Gordon Grant, and Tom Spies spent most of May and June helping "hammer out" a Timber/Old-Growth /Spotted-Owl/Salmon Forest Management Plan for review by President Clinton and other Cabinet members. This effort put tremendous pressure on the GIS shop, and pulled resources away from on-going LTER projects. The Andrews LTER was well positioned to handle this additional workload. The Forest Conference and subsequent Forest Plan development was an excellent forum to demonstrate the LTER's scientific and information management expertise in wrestling with a major societal problem. This is somewhat analogous to the work the Sevilleta LTER has absorbed in the study of the potentially fatal hanta virus.

An experimental Gopher has been established for GIS coverage metadata, with WAIS keyword indexing. The Gopher can be accessed through the LTERnet Gopher.

A network handbook has been developed for users of the Forest Science network. The handbook serves as an introductory guide for new users, as well as a reference for all users.

The GPS campaign was successful in establishing permanent monuments around the Andrews. Maximum replication errors were generally less than 6 cm for latitude and longitude. Pathfinders are now being used from the permanent monuments for positioning of key study sites.

Also of note was the opening of the 5000-square-foot laboratory at the Andrews field site. The lab was funded through congressionally appropriated dollars to the Forest Service.

5.1.2 ARC (Dave Jones)

The Arctic LTER site in Toolik Lake, Alaska has added a land-water interactions group to the lakes, streams and terrestrial work groups. Support for GIS work in all four areas is being put into place. We have purchased a ten-copy license for the Idrisi (PC) software to enable investigators to make use of these systems on their own campuses. Facilities to make use of our digitizer and to convert between the Idrisi and Arc/Info formats are being put in place at our central computing center. We have upgraded our PC-NFS software from version 3.5 to 4.0 on all machines. This year's field season is in full swing, with sampling planned to add to our baseline of long-term lake, river, and tundra vegetation manipulations. One nice early result is that nitrogen isotope added as a tracer to the Kuparuk River is still detectable 2 years later. Use of GIS, isotope and process models to synthesize, interpret, and scale up the results to the whole slope (north of the Brooks Range) is making progress. AVHRR (twice monthly composites) has been obtained for 1990, 1991, and 1992 to support this effort along with other datasets at this scale. Interaction among members of the project via text tools on the Internet has increased. We hope to continue this growth and to add visualizations with the facilities of the GIS on a regular basis.

5.1.3 BNZ (Mark Klingensmith)

The Bonanza Creek Experimental Forest (BCEF) was hit this last winter with a near record snowfall that also came early in the season. As a result we have spent a considerable amount of time this field season repairing snow damage to field equipment, litter traps, and drought shelters. We are now pretty much back in shape and ready for more snow this coming winter!

We are now at the close of our first GPS campaign using the LTER network hardware. Our objectives for this GPS survey included bringing high-level control into the Experimental Forest at easily accessible locations for future surveying efforts, providing ground control points for rectifying and georeferencing GIS layers, and developing a GIS base map of roads and waterways.

Using existing control points we have established 12 monuments scattered throughout BCEF; 5 on the floodplain with access by river, and 7 in the upland on the road system. These were marked with 3' aluminum monuments with flared ends to prevent frost heaving.

Although we had a considerable amount of help from Jeff Jefferson, the arrival of the GPS hardware and software didn't go very smoothly (no doubt as the result of personnel changes at NIN).

The LTER network certainly needs to discuss this situation and make sure that in the future we continue to have a system for coordinating access by LTER sites to the GPS hardware and software.

This past year I have been part of a working group with the UAF Center for Global Change and Arctic System Research developing a directory of arctic and sub-arctic datasets available throughout the Fairbanks campus. I have put the draft copy of the directory into a local WAIServer for experimental access across campus. At this point I need more experience with PC and MAC WAIS clients as more prospective users have these platforms than have Unix that I am running on. Having seen the utility of the WAIS software, I would like to expand the WAIServer to include LTER metadata.

We are using supplemental money to upgrade our optical jukebox to a system that uses standard Unix file systems rather than proprietary data formats on the disks. I felt that this was important for future availability of archived data and because the new system will use more standard operating system utilities for jukebox functions making the whole system easier to administer and more reliable. The jukebox drive will also use high-density optical disks and have a larger total file system capacity.

5.1.4 CDR (Abderrahman El Haddi)

Summer 1993 turned out to be another excessively rainy, cold, and busy year. We have recorded some very low temperatures during the spring and summer seasons. I am expecting to see more than double the normal precipitation, and less than half the normal growing degree days for the 1993 growing season. This is imposing abnormal schedules, since we need to look at the data as it is generated, and look for patterns. This is being done successfully, because this summer we achieved over 98% electronic data collection. The high degree of automation was achieved with the addition of five HP 100LX palmtops and Eclaire software to our data collection and mobile computing "arsenal". All data about plant species percent cover, counts, number of Gophers, etc., are now entered and checked in the field as they are collected. This translated directly into an improvement in the quality of the collected data, as measured by number of errors found, and led to a reduced cost and immediate availability of the data for analysis.

5.1.5 CPR (Tom Kirchner)

Work has continued over the last year on the Distributed Process Interface Library that is being developed to support the Intersite Climate Database project. Documentation is now being prepared, and ltermenu, which makes use of the DPIL, is being ported to PCs under DOS. Our metadata files have been converted to use the extensible format described at last year's meeting. Ltermenu is currently being converted to access information from the new metadata format.

The bibliographic routines developed at Central Plains Experimental Range (CPER) have been converted from FORTRAN to C++, and support has been provided to generate FrameMaker documents as well as ASCII files. Further development is planned to create filters for generating the reference section for papers from the citations within the papers. Internet access to the CPER bibliography will also be provided using a tool developed from the DPIL. We are also investigating the generation of hypertext links between the FrameMaker bibliography file and the CPER database that will allow one to view the bibliography in FrameMaker, select a reference, then examine the data and associated metadata used for creating the paper. We are also investigating creating hypertext links between the database and the bibliography system.

5.1.6 CWT (Gil Calabria)

During the past year, considerable energy has been spent by the data management personnel in making data access and visualization an integral part of the Coweeta LTER Database System. From these efforts, three on-going projects were initiated.

First, all of the Coweeta Bibliographic References, Data Bank Catalog, Core Data Sets Catalog, On-Going Research, Personnel Directory, and other relevant site information were made available on-line with the creation of the Coweeta Gopher Server. This service allows any user to either browse a hierarchy of documents, or search for documents that contain certain keywords or phrases.

Second, since the interface between Ingres databases and the SAS package is not to be released until the end of the year, all the Coweeta LTER datasets have been converted to SAS datasets. By using the SAS-Assist Module, any investigator is capable of interactively graphing, analyzing, and summarizing any archived dataset.

Thirdly, initial thought has been given to the design and implementation of a scientific relational database management system by using the upcoming SAS-Access and ESRI's ARC/Ingres modules to interface with Ingres databases. This will allow us to take advantage of a relational scheme for data access, link the data with its respective georeferenced information, and use statistical modules to interpret the dataset.

Finally, since most PIs and their respective graduate students are not familiar with the Unix environment, we have recently purchased X-Emulation software to further improve system access and develop a more user-friendly interface to all the services provided by the Coweeta DM & GIS Lab. Because of the anticipated workload generated by this X-Interface, we have proposed to use technical funds from NSF to purchase a true server-level workstation from Sun Microsystems. In addition, we are also purchasing two PC-486s and a Mac computer for three PIs who currently have older model machines which cannot handle the proposed X-Emulation software.

5.1.7 HBR (Cindy Veen)

We are working with the University of New Hampshire to put the data we now have available on our public access bulletin board (Source of the Brook) on the Internet. It has been a slow process, but we hope to have it ready sometime this fall.

We continue to slowly add datasets to our archive and public access systems, and collections of physical samples to our Sample Archive. The Hubbard Brook publication list is now available on our public access bulletin board. Subsampling of archived samples is beginning to increase; studies have been initiated that potentially could include subsampling of several thousand archived samples. We have revised the procedure for requesting permission to subsample to include submission of a formal subsampling proposal form.

5.1.8 HFR (Richard Lent)

Expansion of hardware and software systems is continuing at the Harvard Forest. Ten new 486-based IBM PCs were recently purchased, running MS-DOS 6.0 and Windows 3.1. Five of the machines also run OS/2. Most of our software packages were upgraded, including database, spreadsheet, statistics, graphics, and word processing/desktop publishing. New software acquisitions include a Windows-based package for geographic coordinate conversions (Geographic Calculator), some new Idrisi GIS modules for image display and conversion of Arc/Info layers, and Windows versions of DOS programs (e.g., Systat, Borland Pascal). Additional computer acquisitions will include contouring and cartographic software for high-resolution output, logistic regression, resampling statistics, and structural modeling programs, and assorted utilities for graphical display and GIS-based data processing and simulation. Planned hardware purchases for this year will include new color output devices (pen plotter, ink-jet printer), a scanner, and high-capacity tape drives to be used in data archiving.

In conjunction with these hardware and software enhancements, work is progressing on the Harvard Forest data catalog and archive. The on-line data catalog will be linked to disk- and tape-based archives of long-term datasets and accompanying documentation. We are also exploring a Windows-based image management system (HyperPhoto) for documentation and archiving of GIS images. In addition, the Harvard Forest library is updating its holdings and computerizing its card catalog and reprint collections. Bibliographic software (Papyrus 7.0) has been acquired for use both by the library and by individual researchers. We also plan on purchasing a bibliographic service, such as Current Contents on Diskette, to provide our field station with current data on the research literature.

A 100-m grid system is being completed this summer over the entire Prospect Hill tract of Harvard Forest, an area of approximately 900 acres. This grid is being installed by a certified professional land surveyor and will be a permanent, highly accurate spatial reference system for mapping, sampling, and other field work. The grid will also provide accurate ground control points for registering remotely-sensed imagery for Harvard Forest.

Harvard Forest has obtained LTER supplemental funding to expand its systematics collections, including both plant and insect taxa. The collections will serve as a repository for voucher specimens collected on the Forest as well as being a spatially-referenced collection for research on the effects of landscape heterogeneity on biological populations. Two summer students, participants in the first summer of our new Research Experiences for Undergraduates program, are working on the collections in conjunction with their work with staff ecologists in the areas of plant and animal population biology.

5.1.9 JRN (Barbara Nolen)

Our office has recently moved from Room 117 to Room 106 in Foster Hall, New Mexico State University. We have purchased two new Gateway 486/66 computers along with a dual speed Toshiba CD-ROM and a Pinnacle Micro Magnetic/Optical drive for back-up. We have the computers networked and are currently setting up peer-to-peer networking between the computers.

We enjoyed hosting the Coordinating Committee Meeting at New Mexico State University in April.

Through the winter we did an extensive GPS project, collecting positions for more than 300 research sites, rain gages, watering points. We also added 24 high-precision positions to our network across the basin.

This was a cooperative effort between the Jornada LTER, US Forest Service, USDA Jornada Experimental Range and the Bureau of Land Management. The data have been processed and will be included in the GIS database.

We are getting our GIS and remote sensing databases archived and on-line. We have a digital elevation model for the Jornada Basin that combines 22 topographic quadrangles of USGS 1:24000 DEMs.

Our site bibliography has been updated and incorporated into the network bibliography.

John Anderson has been diligently customizing and upgrading our climate data from RDL-SIMS to RDL-Master.

5.1.10 KBS (Lolita Krievs)

KBS GIS is acquiring new aerial photos for the KBS properties. Abrams Aerial Survey of Lansing, MI, was contracted to do the flyover and prepare B/W contact prints of the site. The flight took place 7 May 1993; enlargements have been prepared. Arrangements have also been made with Abrams for annual midsummer flights using IR film the first of which was flown 16 July 1993. I will be using the B/W aerials to update our area basemaps while the IRs will serve to supplement our vegetation database. Arc/Info version 6.1.1 software will be used for digitizing and mapping.

KBS is also currently operating two weather stations. Campbell CR10 dataloggers are used to record measures such as precipitation, soil and air temperatures, relative humidity, wind speed/direction, and solar radiation. Procedures are in place which allow daily means from both stations to be automatically distributed via email to interested parties. Comparing data from both stations allow us to more readily screen the values for possible instrument malfunction. I am reviewing the climate data from both stations to determine the acceptable level of variation between and within the two sites. I will use the results to define a series of flags which will allow the user to assess data credibility.

Other ongoing data management activities at KBS include the design and development of an on-line KBS LTER Data Yearbook which will allow current and future investigators to readily retrieve data, descriptors, summary statistics, and/or reference information on any of the core sampling variates. I have prepared a hardcopy working draft of summary statistics and graphs for 17 of these variates. Summary statistics and graphs provide a measure of data quality control.

By evaluating procedures used to create the working draft, I hope to establish links between data files and descriptors which when called will define the parameters necessary for automated standardized statistical output. Automation will in turn provide for consistency and continuity in data format. By evaluating procedures used to create the working draft, I will be able to establish links between data files and data descriptors. These links will provide both a preliminary statistical analysis of the data and serve as the first step in the QA/QC process.

5.1.11 KNZ (John Briggs)

The Konza LTER had their NSF site review during the month of May. John Porter (VCR) was on the site review team and we used this opportunity to update our LTER dataset catalog and our detailed methods manual. At this time, we have not received our formal reviews back, but the oral feedback was positive.

To relieve the time and effort it now takes to handle data requests, (we are averaging ~two per week), we have installed/updated our interactive data access system. This has been done in two ways: 1) updating the Oracle system (KIMS;Konza Information Management System) which uses SQL to access some of our database, and 2) installing a Gopher. (It resides on bison.konza.ksu.edu). These two data access tools now handle about 90% of our data requests.

We are using our technology supplement to: 1) upgrade our Sun (from a 4/110 to a Sun 10) including more disk space to allow more of our database to be on-line using the tools described above, 2) purchase an organic carbon analyzer that will allow rapid, reliable measurements of organic, and inorganic carbon concentrations in liquid samples, 3) update our TDR (time domain reflectometry) system to measure soil moisture, and 4)install a high-speed modem on our Novell network. Thus, we will have three means for non-KNZ scientists to access our database.

We are currently in the middle of an event that occurs (we hope) only once in our lifetime---excess rainfall, which has resulted in two 100-year flood events in 6 months. We have altered some of our sampling regime to take advantage of this event.

5.1.12 LUQ (Eda Melendez)

Since the last LTER Data Managers' Meeting in Waikiki, data management personnel at LUQ have been involved in activities that permit us to achieve two goals: (1) the acquisition of long-term data which can complement our locally gathered datasets, and (2) increased involvement in publication activities. For the former we have obtained, modified (to a more favorable format), and archived long-term data from different sources such as USGS (streamflow, sediment, met station data), NADP (chemistry and rainfall data), and NOAA (rainfall and temperature data). Graphical representations of some of our long-term data were generated for inclusion in our site's annual report. One of our investigators has already used the acquired streamflow data together with our long-term rainfall data to study evapotranspiration at our site. He will present his findings at the ESA Meeting.

For the second goal, we are still in the preliminary stage. We are attempting to identify the publication arenas, as well as determining what constitutes an "acceptable" publication. In our four years we have produced many reports, manuals, guidelines, but we have a lot to learn about this business. As an effort to begin getting involved, we included an article in the most recent DataBits with the purpose of sharing feedback with other data managers on our main topic in our last annual report. A figure depicted the structure we have developed over four years for data management at our site. One of our major accomplishments has been the definition and communication of our data management philosophy. Our site is now preparing the renewal proposal. We are making every effort to communicate our goals of 1) strengthening our routine data management practices, and 2) increasing our involvement at the network level.

5.1.13 MCM (Jordan Hastings)

Robert Wharton, Jr., the PI for the new McMurdo site in Antarctica, attended the opening session of the meeting, and presented an overview of the project. The so-called Dry Valleys, about 100 miles west of McMurdo, are effectively a frozen desert: no rain has fallen in over 2 million years, and average temperatures are below freezing even in Austral summer. Nonetheless, simple life forms exist: nematodes in the thin, rocky soils, and algal mats in some of the lakes, which are perennially ice-covered but recharged by brief glacial melt. Research will concentrate initially on geomorphic controls over the trophic structure of these biota. Access to the field site is only possible during October through February, via plane from New Zealand, approximately 2000 miles to the north.

The seven co-PIs for the McMurdo site are spread out across the United States; all are connected to the Internet at their respective institutions. In addition, there is currently a fractional T1 Internet connection to the McMurdo base facility; this should be upgraded to full T1 capacity within the next year or two. Several dozen micro- and mini-computers are installed at McMurdo, to support the logistics and scientific programs managed there.

Jordan Hastings, the data manager, intends to make a “micro-GIS” of the Dry Valleys the focus of data management activity for McMurdo. This approach has a “sweet” side -- the intuitive appeal of geographically accessed and displayed information -- and a “sour” side -- dealing at the outset with issues of scaling, time series, and multi-dimensional data types (points, lines, spatial, and volumetric measurements) which are confronted in mapping. The micro-GIS will be implemented in Arc/Info on Sun computers at the project’s home institution, the Desert Research Institute; all the co-PIs will have access to this through Arc/View running on MACs and/or PCs, connected across the Internet. Base map data will be loaded during the autumn, and the first experimental data are expected next spring.

5.1.14 NET (Rudolf Nottrott)

The LTERnet Information Server (using the Gopher/WAIS software) has received several new sections, such as WAIS-indexed versions of the Core Data Set Catalog, and the LTER Personnel Directory. Also, the prototype for the All-Site Bibliography, with entries from 15 sites, is accessible there. The prototype bibliography was developed by Caroline Bledsoe and Harvey Chinn, with the help of data managers at most sites. Other sections contain newsletters (PostScript and other formats) and images. Images are in GIF format and can be displayed with a single keystroke (if client software is capable of doing so -- X clients and Mac clients are).

The LTERnet Gopher has been “officially” listed at the central Gopher Internet site (boombox.microu.umn.edu).

We have widened Internet connectivity of the LTERnet system at both high and low speed through the addition of two new access channels -- a 100 MBit/sec FDDI interface and a 9.6 KBit/sec Serial Line IP (SLIP) interface.

LTERnet now has a direct link via fiber optic cable (FDDI) to the NSFnet T3-backbone, and has been upgraded from a SPARC 2 to a SPARC 10 workstation. The 100MBit/sec FDDI link has a separate IP number (192.42.145.13) and host name (LTERnet-T3), which resides next to the venerable ethernet interface (LTERnet, 10MBit/sec) on the SPARC 10 workstation. LTERnet, with its functions as a communication and information server, hosts the archive for LTER site satellite images on a 10GByte rewritable optical jukebox. With the need to transfer image files of several hundred Megabytes, the 10MBit/sec ethernet interface presented severe limits when compared with the 45MBit/sec speed of the NSFnet T3 backbone.

The hardware upgrade of LTERnet to SPARC 10 required a software upgrade for the optical jukebox. Unfortunately, the jukebox upgrade has not proceeded smoothly, and we still have intermittent problems writing to the archive (retrieval is no problem). We have been working with the San Diego Supercomputer Center to use their Terabyte tape silo for archival storage of LTER data (remotely sensed images only), which would be accessible on LTERnet through its FDDI<->T3 connection. We are also exploring possibilities for similar archival storage on a tape silo at the University of Washington.

The other main computer of the LTER Network Support System, Space, has also been upgraded to a SPARC 10 machine. These upgrades have removed bottlenecks in CPU speed and data transfer speed. (You can try LTERnet's high-speed interface by connecting to LTERnet-T3.LTERnet.washington.edu.)

At the slow end, the SLIP interface is somewhat more complex to set up than a "plain" communications software, such as Procomm, but brings Internet functions to your PC over the phone line. Using SLIP, software packages for e-mail, FTP, and Gopher put messages or files directly onto your PC. For more details on LTERnet's SLIP interface see DataBits, Summer 1993 edition.

5.1.15 NIN (Jeff Jefferson)

Our renewal proposal and subsequent appeal were both denied by NSF. North Inlet will no longer be an official LTER site after October 1994.

We lost two of our PIs (Bill Michener and Liz Blood) and our data manager (Scott Chapal) to the Joseph W. Jones Ecological Research Center in Newton, GA. Jim Morris has taken over as head PI for NIN and Jeff Jefferson has become the data manager. Dwayne Porter, formerly with the South Carolina Marine Resources Division, has been hired as our GIS Lab Manager.

We are in the process of archiving our complete LTER dataset. We hope to have this archived with the Network Office by the All Scientists Meeting.

Our new field lab was dedicated on 8 May 1993. This new facility was built with FEMA funds granted following the destruction of the old lab by Hurricane Hugo. This new facility has full ethernet connectivity. Our site is now fully connected.

We managed to get a no-cost extension to a LTER supplement that had mistakenly been terminated with the old LTER grant. This enables us to fund a GPS workshop to be held following the All Scientists Meeting in September. The workshop is designed to familiarize all the sites with the new GPSurvey post-processing software that was also obtained with the supplement.

5.1.16 NTL(Barbara Benson)

This year we have purchased an Oracle Server for the file server on our LAN (Novell Network). We also acquired Oracle Client Tools for the Macintosh. We are in the process of moving our databases into the Oracle client/server environment.

We welcomed a new staff member at the Center for Limnology last fall, Paul Hanson. He is serving the vital role of network administrator. He has developed a menu-driven user interface for the DOS portion of the file server.

We decided to put our bibliographic information into an EndNote Plus database. This database can be accessed by either a PC or Macintosh user.

In our 1993 supplement we requested a Sun SPARC 10 computer for our LAN at the enter for Limnology. We plan to upgrade our Oracle software to the Sun platform later this year.

The water level in Lake Mendota was 26" over normal summer maximum at its peak value this summer.

5.1.17 NWT (Rick Ingersoll)

We have made a number of important software acquisitions for our Sun SPARC 2 during the past 6 months: FrameMaker, Asterx, and Ingres/ABF. Hardware acquisitions that will upgrade our data management and modelling laboratory include an additional 16-MB memory and 2.3-GB storage for our Sun SPARC 2, as well as 4 X-terminals. Our GIS laboratory has recently acquired SPlus statistical software and a 1-GB optical disk drive.

We have made a great deal of progress in terms of our transition from an investigator-oriented to a centralized data management system. Two students (full-time in the summer and half-time during the academic year) assist our data manager in the manual entry and processing of data. We meet with investigators and graduate students to design their field and laboratory data collection forms, write EasyEntry programs for

data entry, and write shell scripts for additional processing and archiving of the data. All of our procedures and programs are fully documented in two 3-ring binders that are constantly being updated.

Metadata for all NWT LTER datasets are now stored in an Ingres database. Moreover, we now have nearly 40 datasets archived in ASCII format with the metadata. We have used a modified version of the format described in: Conley, W., and J.W. Brunt. 1991. An institute for theoretical ecology? Part V. Practical data management for cross-site analysis and synthesis of ecological information. Coenoses 6.

5.1.18 PAL (Karen Baker)

The plans for the August '93 through March '94 field season are unfolding with continuation of the oceanographic, krill, and bird surveys. In addition to the weekly on station sampling out of Palmer, there are major ship cruises planned off the west coast of the Antarctic Peninsula. Having completed lterjan93 and ltermar93 cruises, a spring cruise in Aug93 and an annual cruise in Jan94 are planned. This completes an intense year of sampling. Development of an experimental Gopher site to provide on-line documentation to investigators continues. Two Unix workstations are being tested at sea in an attempt to provide more real-time analysis, more uniformity between home and field operations, as well as to make our historical documentation and data available in the field.

5.1.19 SEV (Troy Maddux)

A number of personnel shifts have recently occurred. Bruce Milne has become the new Principal Investigator replacing Jim Gosz who is now at NSF as the Division Director of Environmental Biology. Greg Shore, formerly database coordinator, is now a full time GIS analyst. Troy Maddux, formerly head plant technician, is now the vegetation studies program manager.

A lot of effort has been put into remote sensing this year following the NASA workshop and SAR missions. New hardware and software for dealing with remotely sensed imagery has been installed including 3 Gb of hard disk on the server and a 5-seat IMAGINE license.

The T1 connection to the field station is now fully functional, most of the bugs have now been worked out, one year after installation. With this installation the 7 meteorological stations are now being collected via radio and placed directly on the server using Campbell and NFS software. Desqview is being tested to allow for real-time monitoring of met station information from the data offices on campus.

Short takes:

The recent Hantavirus outbreak has activated a lot of rodent analyses and provided a good opportunity to demonstrate the benefits of good data management.

Papers from the EIM symposium are being offered for review from a WAIServer on the Sevilleta system.

A massive vegetation mapping effort is under way involving the US Fish and Wildlife Service and The Nature Conservancy heritage program.

Everyone is getting geared up to host the 5 Chinese data managers in September.

5.1.20 VCR (John Porter)

The VCR has made a major move towards using electronic services to tie together an increasingly diverse and far-flung PI community. Specifically we have implemented daily and weekly electronic mail calendars. These calendars include information on activities (meetings, trips to the site, boat availability) and are automatically sent to PIs and students. Additionally, we have started operating a Gopher electronic information server (host: atlantic.evsc.Virginia.EDU) which is nationally accessible. The Virginia Coast Reserve Information System (VCRIS) includes access to current weather conditions, publication lists and bibliographies, widely used data (such as tides and weather) and administrative documents (proposals etc.).

We are also working on increasing our ability to collect more types of automated data in an electronic form. We are working on automating our network of analog ground-water gauges. We have also acquired a total survey station with a survey computer that permits immediate downloads of survey data. With this station it is possible to collect data at a rate at least an order of magnitude faster than with our old manual survey equipment.

We recently completed a successful collaboration with NASA on the SCAR-A campaign. We hosted a sun photometer (a small robot that takes readings on illumination at specific distances from the sun and can be used to derive atmospheric particulate levels) and provided supplemental radiation data. Our role as a collaborator on the project will give us early access to the remote-sensing data acquired for the project, which includes a couple of TM scenes, numerous AVIRIS and MODIS overpasses and high-resolution aerial photography.

5.2 Workshop Agenda

Thursday, 7/29:Arrivals

7:00-9:00 pm Welcome Reception

Introductions and Site Flashes

Friday, 7/30:

7:15 am Breakfast

8:00 am Old Business

- New Sites, guests
- publication arenas for DM-type papers
- ERIM curricula
- All Scientists' Meeting presentation and demos
- hands-on training in data exchange techniques
- common format for data exchange
- relationships with professional societies
- CERN collaboration (Brunt)
- Expansion of LTER Core Data Set Catalog

9:00 am New Business

- NASA Collaboration
- Risser LTER Review Report (Briggs)

10:00-10:15 am Break, Refreshments

10:15 - On-line Information Services Using Gopher and Wais,
(Porter, Nottrott)

-Information Server on LTERnet, Gopher/WAIS

- All site bibliography project.

11:00 - Site Data Management Policies Discussion.

12:00-1:00 pm Lunch

Friday Afternoon

1:00 - Jim Gosz, Division Director of Environmental Biology at NSF on data management and networking.

3:00-3:15 pm Break, Refreshments

3:15 - Working Groups Breakout

4:00 - Tour and demo at Environmental Remote Sensing Center at the University of Wisconsin.

7:00-... Dinner (Midwestern Food)

Saturday, 7/31:

7:15 am Breakfast

8:00 am Meeting begins

8:15 - Guest Speaker Blanche W. Meeson, Project Manager for the Pilot Land-Data System (PLDS), NASA Goddard Space Flight Center.

Blanche will give an overview of her work, with special emphasis on management of spatial data (metadata, exchange standards, integration into scientific information systems, data access and archiving).

9:00 Discussion

- on subjects presented by Blanche, with a view of finding areas in which collaboration is desirable and feasible,

- and Management of Spatial Data. What metadata is needed? How should it best be stored and queried?

10:00-10:15 am Break, Refreshments

10:15 - Working/Discussion Groups - Old and New TBA

12:00-1:00 pm Lunch

1:00 - General Discussion on Working Group Topics

2:00 - Working Group Enhancements / Report Production

3:00-3:15 pm Break, Refreshments

3:15 -Software Demos / Working Groups if Necessary

4:30 pm Wrap up - Action Items Identified

5:00 pm Meeting Adjourns

Sunday, 8/1: ESA Meeting or Departures

5.3 List of Data Management Workshop Participants

Note: Up-to-date address information is maintained on the LTERnet Information System personnel database. For instructions on access to the system, send any message to Overview@LTERnet.edu

Susan Stafford (AND)

Oregon State University, Department of Forest Science,

Forestry Sciences Lab 020, Corvallis OR 97331-7501

LTERnet address: SStafford@LTERnet.edu (or SStaffor), E-Mail: stafford@fsl.orst.edu

Phone: (503) 737-6581, FAX: (503) 737-1393

Donald L. Henshaw (AND)

USDA Forest Service, Pacific NW Station,

3200 SW Jefferson Way, Corvallis OR 97331

LTERnet address: DHenshaw@LTERnet.edu, E-Mail: henshaw@fsl.orst.edu

Phone: (503) 750-7335, FAX: (503) 750-7329

David Jones (ARC)

Marine Biological Laboratory, Ecosystems Center, Woods Hole MA 02543

LTERnet address: DJones@LTERnet.edu, E-Mail: djones@lupine.mbl.edu

Phone: (508) 548-3705, FAX: (508) 457-1548

Mark Klingensmith (BNZ)

University of Alaska, Forest Soils Laboratory,

305 O'Neill Building, Fairbanks AK 99775-0740

LTERnet address: MKlingensmith@LTERnet.edu (or MKlingen),
E-Mail: mklingen@taiga.lter.alaska.edu, fnmlk@alaska (Bitnet)

Phone: (907) 474-7036, FAX: (907) 474-7439

Thomas B. Kirchner (CPR)

Colorado State University, Natural Resource Ecology Laboratory and Department of Range Science,

Fort Collins CO 80523

LTERnet address: TKirchner@LTERnet.edu (or TKirchne), E-Mail: tom@chloris.nrel.colostate.edu

Phone: (303) 491-1986, FAX: (303) 491-1965

Gil N. Calabria (CWT)

University of Georgia, Institute of Ecology, Athens GA 30602

LTERnet address: GCalabria@LTERnet.edu (or GCalabri),
E-Mail: gil@sparc.ecology.uga.edu, gcalabri@uga (Bitnet)

Phone: (706) 542-5691, (706) 542-2968, FAX: (706) 542-6040

Cindy Veen (HBR)

USDA Forest Service, Forestry Sciences Laboratory

P.O. Box 640, Durham NH 03824

LTERnet address: CVeen@LTERnet.edu, E-Mail: c_Veen@unhh.unh.edu, or S24L06A (USFS)

Phone: (603) 868-7643, FAX: (603) 868-1538

Richard A. Lent (HFR)

Harvard University, Harvard Forest, PO Box 68, Petersham MA 01366

LTERnet address: RLent@LTERnet.edu, E-Mail: harvard.forest (Omnet)

Phone: (508) 724-3302, FAX: (508) 724-3595

Barbara Nolen (JRN)

New Mexico State University, Department of Biology,

Box 30001, Dept. 3AF, Las Cruces NM 88003

LTERnet address: BNolen@LTERnet.edu, E-Mail: bNolen@nmsu.edu

Phone: (505) 646-4465, FAX: (505) 646-5665

Lolita Krievs (KBS)

Michigan State University

W.K. Kellogg Biological Station

Hickory Corners MI 49060-9516

LTERnet address: LKrievs@LTERnet.edu, E-Mail: krievs@msukbs (Bitnet)

Phone: (616) 671-2214, FAX: (616) 671-2351

John M. Briggs (KNZ)

Kansas State University, Division of Biology, Ackert Hall, Manhattan KS 66506-4901

LTERnet address: JBriggs@LTERnet.edu, E-Mail: jmb@andro.konza.ksu.edu

Phone: (913) 532-6629, FAX: (913) 532-6653

Eda C. Melendez (LUQ)

Terrestrial Ecology Division, P.O. Box 363682, San Juan PR 00936

LTERnet address: EMelendez@LTERnet.edu (or EMelende),
E-Mail: e_melendez@upr1.upr.clu.edu

Phone: (809) 767-0350, FAX: (809) 758-0815

Jordan T. Hastings (MCM)

Biological Sciences Center, Desert Research Institute, P.O. Box 60220, Reno, NV 89506

LTERnet address: JHastings@LTERnet.edu, E-Mail: jordan@maxey.unr.edu

Phone: (702) 673-7445, FAX: (702) 673-7397

Rudolf W. Nottrott (NET)

University of Washington, College of Forest Resources, AR-10, Seattle WA 98195

LTERnet address: RNottrott@LTERnet.edu (or RNott), E-Mail: rNott@lternet.edu

Phone: (206) 543-8492, FAX: (206) 685-0790

Caroline S. Bledsoe (NET)

University of California-Davis, Department of Land, Air and Water Resources / Hoagland Hall

Davis CA 95616

LTERnet address: CBledsoe@LTERnet.edu, E-Mail: csBledsoe@ucdavis.edu

Phone: (916) 752-0388, FAX: (916) 752-1552

Jeff Jefferson (NIN)

Baruch Institute, University of South Carolina, Columbia, SC 29208

LTERnet address: JJefferson@LTERnet.edu, E-Mail: jeff@inlet.geol.scarolina.edu,

Phone: (803) 777-3926

Barbara J. Benson (NTL)

University of Wisconsin-Madison, Center for Limnology

680 N. Park Street, Madison WI 53706

LTERnet address: BBenson@LTERnet.edu, E-Mail: bbenson@macc.wisc.edu

Phone: (608) 262-2573, (608) 262-3088, FAX: (608) 265-2340

Rick Ingersoll (NWT)

University of Colorado, INSTAAR, Campus Box 450, Boulder CO 80309-0450

LTERnet address: RIngersoll@LTERnet.edu (or RIngerso),
E-Mail: ricki@culter.colorado.edu

Phone: (303) 492-4771, (303) 492-0566, FAX: (303) 492-6388

Karen Baker (PAL)

UCSD/SIO, A-018, La Jolla CA 92093-0218

LTERnet address: KBaker@LTERnet.edu, E-Mail: karen@crseo.ucsb.edu

Phone: (619) 534-2350, FAX: (619) 534-2997

James W. Brunt (SEV)

University of New Mexico, Department of Biology, Castetter Hall,

Albuquerque NM 87131-1091

LTERnet address: JBrunt@LTERnet.edu, E-Mail: jBrunt@sevilleta.unm.edu

Phone: (505) 277-9342, (505) 277-9372, FAX: (505) 277-0304

Troy Maddux (SEV)

University of New Mexico, Department of Biology, Albuquerque NM 87131-1091

LTERnet address: TMaddux@LTERnet.edu, E-Mail: tMaddux@sevilleta.unm.edu

Phone: (505) 277-9164

John H. Porter (VCR)

University of Virginia, Department of Environmental Science, Clark Hall

Charlottesville VA 22903

LTERnet address: JHPorter@LTERnet.edu, E-Mail: jhp7e@virginia.edu

Phone: (804) 924-7761, FAX: (804) 982-2137

Blanche Meeson (NASA)

NASA Goddard Flight Center

Code 902.2

Greenbelt, MD 20771

LTERnet address: BMeeson@LTERnet.edu, E-Mail: meeson@pldsg3.gsfc.nasa.gov

Phone: (301) 286-9282

Jack W. Witham (Holt Research Forest)

College of Forest Resources, University of Maine

Old Stage Road,

Box 309

Arrowsic, ME 04530

LTERnet address: JWhitham@LTERnet.edu, E-Mail: rfy395@maine.maine.edu