THE MANAGEMENT OF ELECTRONICALLY COLLECTED DATA WITHIN THE LONG-TERM ECOLOGICAL RESEARCH PROGRAM

Rick Ingersoll and Scott E. Chapal

January 1992

[1994 Update](#)

ACKNOWLEDGEMENTS

We would like to thank Phyllis Adams, John Gorentz, Mark Losleben, Bernie Moller, Mark Williams, and Connie Woodhouse for their contributions to the survey. We would also like to thank those who contributed to the individual site responses.

INTRODUCTION

The rapid evolution of automated sensing equipment and related computer technologies has been concurrent with the Long-Term Ecological Research (LTER) program's increased reliance on electronically gathered data. Most meteorological data are collected, stored, and/or transferred by electronic means at LTER sites. Although these technologies have freed climatologists and data managers from the tediousness of manual data entry, they also have necessitated quality assurance/quality control (QA/QC) protocol modifications in order to achieve acceptable levels of accuracy, precision, and error in the data. Such modifications (1) require identification of additional sources of poor-quality data, and (2) must account for differences or reductions in human involvement at various stages during and subsequent to automated collection of data. We believe it to be an appropriate time for an assessment of the manner in which the network of LTER sites has coped with problems inherent to modern methods of data collection, since inaccuracies resulting from undetected problems could be subtle enough to mask real variability in the data.

The idea for an electronic data collection survey within LTER originated in QA/QC working group sessions at the August 1991 meeting of the LTER data managers in San Antonio, TX. The objectives of the survey were to (1) obtain detailed information concerning the management of electronically collected data within the program, (2) to summarize that information, and (3) to offer recommendations for the improvement of management of electronically collected data within LTER. A more detailed description of the reasoning behind the survey's origins can be found in the explanatory text accompanying the survey (Appendix 1). The survey was prepared in September and distributed to all LTER sites via electronic mail (email) in October. The responses were compiled and this report was prepared during December 1991 and January 1992.

THE RESPONSES

The reply rate was favorable, given the amount of detailed information requested on relatively short notice. Responses were received from 14 of 18 LTER sites. Moreover, responses were received from 2 National Park Service (NPS) sites. Although the original intention was for the survey to be LTER specific, it was decided to treat the 2 NPS sites as LTER sites when it became apparent that the responses to individual portions of the survey were qualitatively within the range of responses from the "true" LTER sites. The following sites participated in the survey:

H.J. Andrews Experimental Forest (AND), Arctic Tundra (ARC), Bonanza Creek Experimental Forest (BNZ), Central Plains Experimental Range (CPR), Coweeta Hydrologic Laboratory (CWT), Hubbard Brook Experimental Forest (HBR), Harvard Forest (HFR), Jornada (JRN), Konza Prairie (KNZ), North Inlet Marsh (NIN), North Temperate Lakes (NTL), Niwot Ridge/Green Lakes Valley (NWT), Sevilleta National Wildlife Refuge (SEV), Virginia Coast Reserve (VCR), Isle Royale National Park (IRP), and Olympic National Park (ONP).

[The response for El Verde Field Station in Puerto Rico (LUQ) was received too late to be included in the text or tables that follow. Nevertheless, a cursory examination indicated that its inclusion would not significantly alter the content of this report and the response is included in the appendices.]

Note that "site" as used throughout this document refers to one of the sites above and not to individual measurement locations within those sites. These measurement locations were often referred to as sites within the individual survey responses but for the purposes of this evaluation were "lumped" together and treated as a single site.

Individual responses were variable both in terms of the detail provided and their formats, although they generally adhered to the specified format. Some of this variability might be attributed to the short notice in conjunction with personnel limitations at the sites. A number of the sites provided supplementary materials not specifically requested. Ten of the responses were received in the form of electronic mail messages, 3 as hard copies only, 2 in word-processing format on floppy diskettes, and 1 was received via binary ftp (file transfer protocol). The individual responses appear in their entirety in Appendices 2-17 and, with the exception of removal of the survey "questions", are unedited. [The LUQ response may be found in Appendix 18.]

SURVEY SUMMARY

Caveats

During the initial stages of the evaluation of the responses, it became apparent to the authors that neither could dedicate the time required to conduct as thorough an evaluation as desired. To obtain answers for each of the numerous minor questions that arose would have required an effort that would have been disproportionately larger than the gain in knowledge. Moreover, the "snapshot view" analogy of Foster and Boose (email message, 24 July 1991) is certainly applicable in this case and we felt it important that the snapshot be displayed before it has faded. Thus, we worked with what we had.

Neither of us professes to be experts on all of the hardware, software, and techniques mentioned in this report. Thus, we acknowledge the definite possibility of misinterpretation or outright error. The individual survey responses and contact people should be consulted for further clarification.

The absence of "x" for a given site in any of the summary tables should not be construed as indication of lack of applicability of the particular item for that site. The number of "x"s applied to a site in a particular table is to some extent proportional to the amount of detail contained in the response for that site.

The mention of, or comments regarding, commercial products does not constitute any official endorsement by any of the parties involved, i.e. the National Science Foundation, the Long-Term Ecological Research program, the National Park Service, or the U.S. Forest Service.

1. Variables

Climatological variables are those most commonly measured via electronic means at the 16 sites (Table 1). Air temperature, wind (speed and direction), and solar radiation are monitored electronically at all but one of the sites. At least 75% of the sites use datalogger-type systems to quantify relative humidity, precipitation, and soil temperature. Thus, modern automated technology forms the backbone of climatological data collection within LTER.

2. Measurement Periods

The reliance on electronically collected climatological data within LTER is a comparatively recent phenomenon. If one sums the number of primary variables measured over all sites this becomes readily apparent. In 1979, only 4 primary variables were measured electronically within LTER. In 1991, the number had reached 135, with most of the increase occurring in the mid-1980s.

3. Measurement Frequencies

Details on measurement frequencies may be found in the individual survey responses (Appendices 2-17).

4. Hardware

Dataloggers manufactured by Campbell Scientific Instruments (CSI) were the most commonly used loggers within LTER (Table 2). Versions of these instruments have been used at 14 of the 16 sites, with the CR21X and the CR10 being the most commonly employed models. The only other makes of automated datalogging hardware that have been used at more than one site are Omnidata and Li-Cor.

Individual site comments are acknowledged parenthetically throughout the remainder of the report.

Advantages of the CSI dataloggers were reported primarily as having greater flexibility and programmability (AND, HFR), and the sites that had reported using these instruments were in general very satisfied (e.g. ONP). Some did not see the need for the additional flexibility provided by the CSI loggers, however, and were more concerned with reliability (BNZ). Advantages of the Li-Cor dataloggers were user-friendliness, retention of data and programs during temporary power losses, low cost, mobility, and durability (IRP, HFR).

The LTER program, when viewed as a whole, has amassed a great deal of experience with hardware available for electronic data collection (Table 3). More than 75% of the sites have used CSI temperature/relative humidity probes and 50% of the sites have used CSI wind speed and direction sensors. Greater than 50% of the sites have used Li-Cor radiometers for measuring solar radiation. Precipitation measurements at 75% of the sites have been made using gauges manufactured by Belfort Instrument Company, Sierra-Misco, and Texas Electronic. Soil moisture has been determined using 227 Delmhorst gypsum blocks at greater than 25% of the sites. CSI data storage modules were used at 50% of the sites and CSI cassette interfaces were used at nearly one third of the sites. It was not clear the extent to which the reliance on CSI for peripheral hardware was determined by usage of CSI dataloggers, but factors such as compatibility certainly were mentioned in the surveys.

Although site-specific problems were evident when climatic environment and electronic data collection systems were considered together, the vast majority of problems at most sites have been experienced at several other sites (Table 4).

Note: Preventative maintenance protocols presented in the survey responses in some cases were interpreted as indicators that the potential for a particular problem existed. Thus the entries in Table 4 do not necessarily mean that a problem actually was experienced at each of the indicated sites, but rather that it either had or was likely to have occurred.

Power reductions (or losses), the effects of extreme temperature, the impacts of animals and insects, the effects of moisture and condensation, and deterioration of equipment and sensors are each problems that have been reported by at least 50% of the sites. Additionally, problems associated with reading of magnetic tapes, measurement of relative humidity, dirty sensors, and data storage losses were each reported by more than one third of the sites. Obviously, there is "overlap" of the general categories since one problem often generates others. Extremely low temperatures, for example, may affect both power supply and the operation of tape recorders. Nevertheless, given personnel changes, the fact that the survey responses were completed in many cases by more than a single individual, and the lack of detail contained in some of the surveys, it is most likely that these are conservative estimates of the extent to which these problems have surfaced within LTER.

Although the most detailed information concerning these problems and their solutions can be found in the individual survey responses, this is an appropriate place to summarize some of the easier solutions to common problems. Power reduction and loss effects can be minimized by greatly exceeding the manufacturer's recommendations for both power reserves and grounding, as well as ensuring that system resistance is kept low (NWT). The effects of low temperature on tape recorder operation have been minimized through the use of thermostatted heat tape (KNZ). Exhalation near tapes and/or tape heads should be avoided during cold weather (AND). Other magnetic tape reading problems have been minimized by elimination of the use of high-quality tapes (less noise) and by matching the heads of the machines used for playback with those used for recording (BNZ). Radio telemetry has been an effective alternative to magnetic tape transfer of data (NWT).

Mammals, and particularly rodents, have chewed on sensors, wires, and other field equipment. Isolation of equipment in PVC conduit (AND) or by burial (BNZ) have been successful tactics. Spider webs have obscured sensor surfaces (JRN), created electrical shorts (SEV), and interfered with the moving parts on some instruments (VCR). Moth balls have been used with success in the latter instance. The effects of moisture/condensation were best avoided through preventative maintenance, e.g. frequent examination of instruments and replacement of dessicant. Enclosure of sensitive instruments (data loggers and terminal blocks) in picnic coolers have worked at HBR.

Deterioration of equipment and sensors likewise was best avoided through preventative maintenance, specifically by the timely replacement of suspect components. Awareness of a problem common to the electronic measurement of relative humidity (RH) in LTER provided the initial impetus for the current project (Appendix 1). This and many of the other RH measurement problems seemed to be common to the Phys-Chemical Research PCRC-11 RH sensor used in CSI temperature/relative humidity probes. There is evidence that these sensors can become temporarily ineffective (VCR) or even permanently damaged (NWT) when saturated. Moreover, their "reliable" lifetime under even normal conditions has been questionable (BNZ, CWT, NWT, SEV, IRP, ONP). While replacement times have been generally on the order of a year, these sensors were being phased out at several sites (BNZ, NWT, VCR). Alternative replacement sensors and/or probes have been identified at BNZ and VCR. Wind measurement equipment has also been prone to relatively rapid deterioration, with the weakest links being in failure of windspeed sensor bearings (CPR, JRN, NTL) and wind vane potentiometers (NTL, NWT, SEV).

Gypsum blocks were ineffective for the measurement of soil moisture under very dry conditions (AND, SEV), as well as in saturated sandy soil (BNZ).

5. Maintenance

The frequency of maintenance depended on the particular procedure and piece of equipment involved, and varied from daily to biennially. Seasonal differences in maintenance frequency were reported at 6 of the sites.

6. Calibration Similarly, the frequency of calibration depended on the particular piece of equipment involved, and varied from weekly to biennially. In general, calibrations were less frequent than maintenance procedures. Only 2 sites reported seasonal differences in calibration frequency. Typically individual sensors were replaced rather than calibrated and large, complex instruments were returned to the manufacturers for calibration.

7. Storage, Transfer, and Archival

Internal memory/solid state were field storage media at more than 50% of the sites (Table 5). Data storage modules or packs also were used at more than 50% of the sites. The latter have minimized data losses through increased storage capacities at a number of sites. Although magnetic tape was used for field storage at only 3 sites, more than one third of the sites employed this medium for transfer of data from the field to the laboratory. Solid state and phone cable were each used at about one third of the sites for data transfer.

All but 2 of the sites reported computer hard disks as archival or storage media (Table 6).

Floppy diskette and magnetic tape were the other preferred media, with usage acknowledged by 13 and 9 sites, respectively. Nearly 50% of the sites reported ASCII as a storage or archival format.

8. Quality Assurance/Quality Control (QA/QC)

The vast majority of the sites employed redundant measurement systems and greater than 50% reported regular examination of equipment and/or data as being components of their field QA/QC protocol (Table 7). A variety of commercial software and programs were used, in addition to a number of "homegrown" programs, for QA/QC in the laboratory. In general, these programs performed range checks and or outlier searches, and often produced graphs which were visually inspected for anomalous values. Files were often manually edited, e.g. questionable values were removed or flagged, to improve the quality of a data set.

A number of sites have, at one time or another, attempted to quantify the relationships among redundant measurement systems. In some cases, concurrence of such systems was limited and represented the transition from one data collection system to another. With the exception of dewpoint temperature, correlations between data collected with Interface Instruments M4 (which contained similar sensors to the previous chart recorder system) and those collected by CSI CR21X dataloggers which superseded it were very high (AND). Correlations between chart recorder and CSI CR10 temperatures were also high (r2 > or = .98) (AND). Although differences among redundant measurement systems have been reported as insignificant (JRN), there have been cases of significant diferrences between datalogger data and those collected by other means (HBR). Adetailed comparison of chart recorder, datalogger, and datapod data was recently carried out at NWT. The results indicated that the relationships were complex and frequently nonlinear. Moreover, there is evidence that it is unclear as to what are acceptable differences between the primary and backup/calibration measurement systems (e.g. CWT).

The ablility to detect problems quickly is a necessary component of an effective QA/QC program and some newer products may be useful in this regard (BNZ).

9. Ideal System

Although a number of sites complained of inadequate precision or accuracy for some measurements (Table 4), it is also apparent that the sensitivity of some of the instruments was greater than required. A better parameter estimate is likely to be achieved with replication, rather than with more precise instrumentation (BNZ). Greater precision is not without cost and at least one site (NWT) is planning to continue use of chart recorders for some primary measurements.

Greater compatiblity among the products produced by the different manufacturers would be desirable (IRP) and would certainly provide greater flexibility in any electronic data collection system.

Photovoltaics (NIN, NWT) were reported as being the ideal power source for remote sites.

Of course, the ideal system would be "something that works all the time and never gets out of 'date'" (KNZ).

10. Contact People

The following names were provided as contact people for clarification of survey responses. Additional information (e.g. addresses, phone numbers, electronic mail addresses) is available in the individual survey responses.

AND Fred Bierlmaier

ARC Jim Laundre

BNZ Phyllis Adams, Bob Schlentner, Mark Klingensmith

CPR Tom Kirchner

CWT Barry Argo, Lloyd Swift, Bob McCollum

HFR Gary Carlton

HBR Cindy Veen

JRN Barbara Nolen

KNZ John M. Briggs

NWT Mark Losleben, Rick Ingersoll

NIN Danny Taylor, Scott E. Chapal

NTL Tim Kratz

SEV Douglas Moore

VCR Dave Krovetz

IRP David Toczydlowski, Robert Stottlemyer

ONP Robert Edmonds, Roger Blew, Tina Lipman

CONCLUSIONS AND RECOMMENDATIONS

During the past decade, the LTER has embraced electronic data collection systems, a young and growing technology, for the measurement of a number of variables. Differences existed among the individual sites in terms of variables measured and hardware, software, and protocols employed. Nevertheless, there was considerable common ground as evidenced by the problems encountered and the solutions employed to deal with those problems.

Climatological variables were measured at virtually every site and, for the most part, these data were being collected by automated electronic equipment. Given the high visibility of the LTER program (and the variety of biomes it represents), increased interest (and opportunity) for data exchange, and rising concern over global climate change, it is likely that the LTER climatological data sets will undergo increasing scrutiny in the coming years. For this reason, and because reliance on a rapidly evolving technology dictates a dynamic situation, it is important that the evaluation of electronic data collection within LTER be considered an ongoing and regular process.

The immediate task is determination of the form of that process. The ideas and suggestions contained in the following paragraphs are meant to provide a starting point for a discussion that ultimately leads to a LTER strategy for the regular evaluation of electronic data collection. The suggestions and options are not necessarily mutually exclusive and some aspect of each might be incorporated into such a strategy.

Individual sites and the LTER program as a whole would benefit from more frequent exchange of information on both the inter- and the intra-site level. It seems very likely that the individuals (field personnel) actually involved in the electronic collection of data at the sites often are insulated from one another and are left to their own devices in terms of purchase decisions and solution of problems. While experimentation, experience, and intuition are satisfactory at dealing with these sorts of things, the knowledge of prior success is generally more expedient. A good first step toward information exchange at this level would be for the appropriate personnel to have access to a complete (i.e. including the individual survey responses) copy of this document as soon as possible. It is therefore recommended that the LTER network distribute a copy of this document to the data manager at each site with instructions to make it known by and available to all field personnel involved in electronic data collection. In addition, a copy should be made available to each of the two National Park Service sites that contributed responses. The formation of a new group, on electronic data collection technology, within the LTERnet mail forwarding system would provide a means of continued exchange of information, once the bulk of this report has become obsolete.

Superimposed on this is another information gap that probably exists to varying degrees at a number of sites. The field person responsible for the collection of the climatological data, the data manager, and the investigator using the data are, in many cases, different individuals. These individuals should be encouraged to interact on a regular, if infrequent, basis at each site to ensure that data are of acceptable quality. The LTER Network Office, Climate Committee, and data managers could provide some leadership here. The Standardized Meteorological Measurements for Long Term Ecological Research Sites prepared by the LTER Climate Committee in 1986 was necessarily general and thus did not specifically address electronic data collection. Now that it has become apparent that the large majority of climatological variables in LTER are being measured electronically, the Standardized Meteorological Measurements for LTER needs to be updated and modified to reflect this awareness. A section on troubleshooting would be beneficial (ARC).

Since technology has generated the need for a new document in 5 years, it is recommended that this new SMM for LTER should be in ASCII format and available online in much the same way that the LTER Personnel Directory is. The Electronic Data Collection Survey could be streamlined and modified so that it could provide much of the information necessary for an annual update of the online SMM for LTER. Checklists (similar to the "items" in the tables in this report) could be provided, for example, to facilitate the completion as well as the evaluation of the survey.

The LTER should become active in the provision of feedback to the manufacturers of electronic data collection hardware and software. The survey responses and other sources indicate that the leaders in this industry are generally responsive to the needs of their customers. Presumably, responsiveness would be enhanced if they were approached by an organized group rather than by an individual from one or another site. A working group could be formed that includes representatives from the scientists, data managers, and field technicians within LTER. This group could communicate on a regular basis via email, be involved with distribution and evaluation of the survey, and meet on an infrequent but regular basis (e.g. at All Scientists Meetings or data managers meetings). Moreover, this group would be in contact with the manufacturing industry for the purposes of providing feedback regarding their products. Such feedback should include needs and suggestions, as well as notification of problems. Additionally, interaction with other scientific organizations such as the Department of Energy Research Parks, National Center for Atmospheric Research, National Oceanic and Atmospheric Administration, National Park Service, National Weather Service, U.S. Geologic Survey, etc. could be mutually beneficial.