In Michener, W.K., J.H. Porter, and S.G. Stafford. 1998. Data and information management in the ecological sciences: a resource guide. LTER Network Office, University of New Mexico, Albuquerque, NM

TECHNOLOGICAL UNDERPINNINGS: SOFTWARE

Karen S. Baker

Scripps Institution of Oceanography, University of California at San Diego,

La Jolla, CA 92093-0218

Abstract. Survey results from Long-Term Ecological Research (LTER) sites provide an overview of the variety of software choices made at individual locations where PC, Macintosh and UNIX platforms predominate. The survey considered software categories including bibliographic, data entry, database management systems, drawing, geographic information systems, graphics, spreadsheets, statistics and word processing. The objective of the survey was to assess the heterogeneity of software used by the LTER community.

INTRODUCTION

Decisions with respect to software selection must balance the often conflicting requirements of addressing immediate local community needs and meeting broader, long-term institutional objectives. Research institutions, including biological field stations, often have very specific needs that require further balances between simple versus complex, individual versus standard, and the current state-of-the-art versus emergent technologies. Although software ideally should be extensible and have clear export paths, there are few other specific rules to guide software infrastructure. Software choices often depend upon several factors:

- What are the computational, data management and storage priorities?
- What software options exist? Do options vary by platform?
- What are the costs of hardware, software, and support?
- Where will the different functions of data collection, management (e.g., entry, processing, archival), and analysis be carried out?
- Who is available for technical support (e.g., local support environment, consultants, vendors)?

Consideration of such questions permits a definition of priorities. Subsequently, more task-specific questions can be asked, such as whether scientific visualization tools are a high priority and how data can be accessed.

Available resources play a significant role in discussions of the organizational approach of a research group. Some sites identify and encourage use of a common set of software tools. In such cases, the availability of training can help those users not familiar with the supported tools. Other sites find a range of diverse software to be advantageous. It is important, given the variety of software available, to consider the collective

consequences of software choice and to develop a policy regarding which packages will be site supported in order to maintain realistic user expectations.

Factors such as cost (including the availability of academic discounts), stability, marketing, interoperability, power and ease of use influence decisions about software. Education helps build consensus, but it is important to recognize how software acquisition is influenced by a diversity of legacies (hardware and software), interfaces (human and hardware), and data volumes (small to large). Familiarity also plays a role in the decision-making process. The LTER software survey permits sites to place their own decisions into a network-wide context, i.e., a survey extends a single site's experience to a network of sites.

LTER SOFTWARE SURVEY (1992 to 1997)

The LTER software survey began in 1992 with eighteen sites in addition to the LTER Network Office. By 1997, the survey included twenty-one participants. Results from an earlier survey have been discussed previously (Porter et al. 1996). A yearly LTER software survey of more than nine categories quantifies the diversity and trends of software within the LTER community. Table 1 summarizes bibliographic, data entry, database management systems, drawing, geographic information system (GIS), graphics, spreadsheets, statistics and word processing software products that were employed in 1997. The first line of each category in the table gives the total number of packages used throughout the LTER Network. This is followed by a list of the software packages with at least three site implementations along with the number of implementations at all the sites by platform type.

Table 1. Software and software summary for LTER sites (minimum of three installations) for 1997.

Original form is found at *http://lternet.edu/im.* The first line gives the total number of packages used by platform followed by lines with the number of sites using a specific package.

	PC	MAC	UNIX		PC	MAC	UNIX
Bibliography*	**7**	**1**	**5**	**Graphics***	**16**	**8**	**7**
Procite™	6	--	--	Excel™	8	7	--
Papyrus™	5	--	--	Deltagraph™	2	3	--
Endnote™	4	4		Lview™	5	--	--
Bibtex™	--	--	3	Matlab™	--	2	2
Data entry*	**11**	**2**	**4**	Quattropro™	8	--	--

Excel™	12	10	--	SAS-graph™	4	--	4
Quattropro™	5	--	--	Sigmaplot™	8	--	--
Lotus™	4	--	--	Slidewrite™	3	--	--
SAS™	--	--	4	Cricketgraph™	--	4	--
Quickbasic™	3	--	--	Framemaker™	--	--	3
Database*	**10**	**0**	**11**	**Spreadsheets***	**6**	**3**	**1**
Access™	8	--	--	Excel™	20	10	--
Dbase™	5	--	--	Lotus™	10	--	--
SQLserver™	3	--	--	Quattropro™	15	--	--
ArcInfo™	--	--	4	**Statistics***	**11**	**9**	**10**
Ingres™	--	--	4	Excel™	16	--	--
Oracle™	--	--	4	Matlab™	1	--	2
Msql™	--	--	3	SAS™	12	--	10
Drawing*	**8**	**8**	**3**	Sigmaplot™	7	--	--
Photoshop™	7	6	--	Splus™	3	--	6
Freehand™	2	1	--	Systat™	10	3	1
Coreldraw™	3	--	1	Cricketgraph™	--	4	--
GIS*	**9**	**1**	**10**	Powerpoint™	2	2	--
ArcInfo™	13	7	18	Statview™	--	3	--
ArcView™	12	--	13	SPSS™	1	--	2
Erdas™	6	--	9	**Wordprocessors***	**5**	**4**	**4**
Erdas-imagine™	4	--	5	Framemaker™	3	2	4
Idrisi™	6	--	2	Word™	18	9	--
Grass™	1	--	7	WordPerfect™	17	2	4

				LaTeX™	--	--	5

Software categories with the highest diversity (number of different packages) across all platforms within the LTER Network included graphics and statistics, whereas the lowest diversity was associated with drawing, wordprocessing and spreadsheet software (Figure 1). Given the differences in hardware, it is interesting to consider distributions by platform type. In general, there were more packages used by the PC than the Mac, except in the case of drawing software. In 1997, the largest variety of packages was related to graphics on the PC. The greatest variety of packages for the UNIX system was associated with database, statistics and GIS categories. Examination of temporal trends from 1992 to 1997 (Figure 2) reveals that the software diversity for some categories increased (e.g., graphics, statistics and drawing) while the diversity for other categories remained relatively unchanged (e.g., word processors and spreadsheets). While UNIX was the platform of choice for most database applications in earlier years, there has been an increase recently in GIS and database software use on both PC and Unix platforms.

Figure 1. The number of different software packages used by LTER sites in 1997 for each platform type.

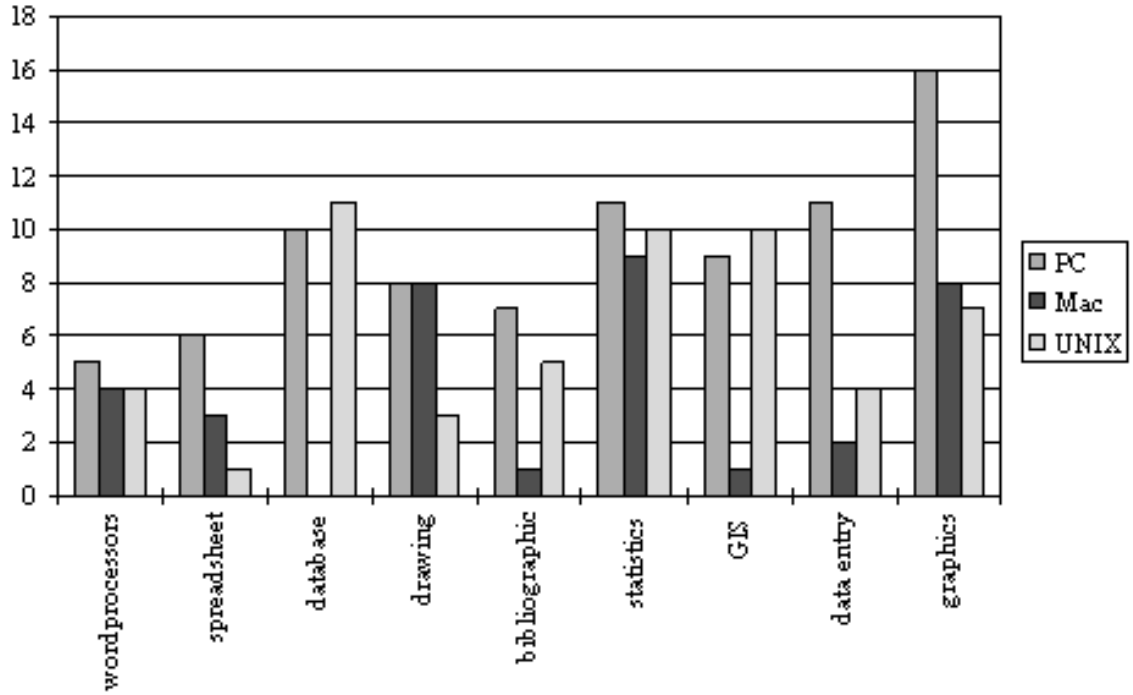
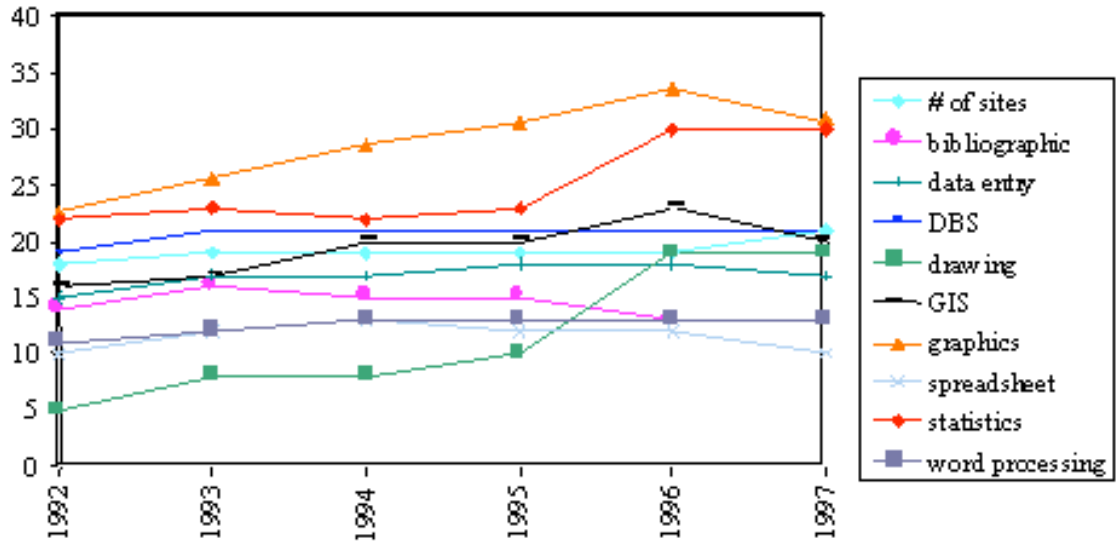


Figure 2. The total number of different software packages (by category) used by LTER sites during 1992-1997.



CONCLUSION

Several issues have and will continue to influence the software infrastructure at long-term environmental and ecological research sites, including scientific objectives, software policies, cost, and standardization. Given the diversity of software available, a general survey facilitates consideration of a wide range of potential solutions. The interplay of objectives and approach is unique to each research site, so decisions with respect to software vary.

ACKNOWLEDGEMENTS

This work was funded by NSF Grant OPP90-11927. This is Palmer LTER Contribution #151.

LITERATURE CITED

Porter, J.H., R.W. Nottrott, and K.S. Baker. 1996. Tools for managing ecological data. Proceedings of the Eco-Informa Workshop, Global Networks for

Environmental Information, November 4-7 1996, Lake Buena Vista, FL, 11:87-92.