Combining these developments with advances in computer networks can lead to electronic submission of manuscripts, grant proposals, and so on.

Networks will become even more important in the very near future. As a supplement to this report I would recommend two articles from a special issue of *Science* on "Frontiers in Computing": the article on workstations (Crecine, 1986) and the one on networks (Jennings et al., 1986).

The next logical step in computing resources, powerful graphics workstations, have undergone cost reductions even between the time that this report was written and the time it was published. In the report the cost of a workstation is given as $20,000 to $40,000. Both Sun Microsystems and Digital Equipment are now selling powerful monochrome graphics workstations with local area network interfaces for around $10,000 with the usual university discounts. When the network becomes sufficiently large to incorporate a file server, the incremental cost of adding a "diskless" node can be as little as $5,000.

Such cost reductions will have ramifications for researchers in both large and small statistics departments, not the least of which is that the real cost of computing will shift even more from the cost of the hardware to the cost of the software and operations. Much statistical software is still sold on a "per CPU" basis. This will represent a severe impediment to a department which may have 20 or more CPUs—it is simply not reasonable to spend $5000 to purchase a workstation only to find that it will cost several thousand dollars a year in software licenses before you can use it effectively. There are two approaches to solving this problem of the cost of software: we must persuade software vendors to issue site wide or department wide licenses and we must rely more on public domain software.

I think that this is an area in which our professional societies can help. With easy access to networks, we can physically facilitate software distribution. For example, Jack Dongarra at Argonne National Laboratories and Eric Grosse at AT&T Bell Laboratories operate "netlib" for numerical analysts. Anyone with access to ARPAnet or USEnet (and possibly others) can send an electronic mail message to netlib requesting the index of available routines or requesting a particular routine. A program running at these sites interprets the message and sends the requested routines back via electronic mail. This is the type of facility that we must work toward because statistical research is a software- and data-intensive activity. Merely having computer hardware is not enough. If we are going to provide facilities for easy network wide distribution of data sets and software, though, we will have to have some form of quality control on the software for it to be helpful. I think this work can best be coordinated through our professional societies.

There are great challenges here, but also great opportunities. We know that the cost of computing hardware is going to decrease to the point that very powerful and sophisticated computing resources are within the budget of any researcher who wishes to take the time to learn how to use them. The report is, again, right on the mark in stressing the importance of standardization of such facilities so they are easily usable. If we can make an effort also to provide, at a reasonable price, the software to enable these facilities to realize their potential, we can look forward to a very exciting and productive time in research.

ADDITIONAL REFERENCES

CRECINE, J. P. (1986). The next generation of personal computers. *Science* **231** 935–943.

JENNINGS, D. M., LANDWEBER, L. H., FUCHS, I. H., FARBER, D. J. and ADRION, W. R. (1986). Computer networking for scientists. *Science* **231** 943–950.

Comment

Edward J. Wegman

I am pleased to have the opportunity to comment on and amplify aspects of this report. The committee that spawned this report had its origins in three-way discussion among Ron Pyke, Ingram Olkin, and me during the 1984 Annual IMS Meeting held in Lake Tahoe, California. It is thus with some avuncular pride that I am able to congratulate Bill Eddy and his colleagues on a job done very well.

In 1983, the Department of Defense began its University Research Instrumentation Program (URIP). As Division Head of the Mathematical Sciences, I saw

Edward J. Wegman is The Bernard J. Dunn Professor of Information Technology and Applied Statistics and Director of the Center for Computational Statistics and Probability, 217 Thompson Hall, George Mason University, Fairfax, Virginia 22030.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *Statistical Science*. www.jstor.org

an interesting opportunity to place compatible computing machines in a variety of mathematical sciences and statistics departments with the idea that similarity of computing resources would create an environment in which easy communication and software and algorithm exchange would be feasible. This idea was discussed among principals at federal granting agencies and thanks to Al Thaler's kind invitation, I was also appointed to the National Science Foundation (NSF) Mathematical Sciences Division's Scientific Computing Research Equipment in the Mathematical Sciences (SCREMS) advisory panel. The Office of Naval Research was the lead agency for the Department of Defense (DoD) URIP in fiscal year 1983.

A goodly number of proposals both to NSF SCREMS and DoD URIP requested either VAX 11/780 or VAX 11/750 computers that first year. Indeed, this computer together with a UNIX operating system quickly became a defacto standard. As technology has matured, we have seen a shift in preferred computer resource. The VAX computers are essentially time-sharing minis which provide computational resources for units such as departments or small research groups. Requested resources in the 1986 and 1987 competitions were much more likely to be workstations such as SUNS, Apollos, Micro VAX, and so on. Fortunately, many of the workstations support UNIX so that communication capability has been maintained even with the shifting hardware technology. It is with this history and the steady accumulation of computing resources by several prominent departments that a documentation of their experiences and perspectives was thought to be useful. For these reasons, I encouraged the IMS to undertake this enterprise while I was still at the Office of Naval Research. In the next couple of paragraphs I would like to make a few observations from the granting agency perspective.

Perhaps the most obvious manifestation of the computer revolution is the ubiquitous microcomputer. It is also perhaps the least significant aspect. The new pipeline and parallel architectures such as systolic arrays, connection machines and hypercubes, the emergence of artificial intelligence, the cheap availability of RAM, the impact of high resolution graphics, the potential for optical, biological, and chemical computing machines, and the pressing need for software/language models for parallel computing are all aspects which have potentially profound implications for statistical science.

In the late 1950s and early 1960s, machines such as the IBM 1620 or IBM 1401 were essentially single-user machines. Users often had direct access and turn around was prompt, if not instantaneous. As machines became larger, more expensive, and more loaded, a despicable monster was created, the University Computer Center. It was filled with employees whose sole role seemed to be aimed at keeping users at arms length by constantly updating, revising, or otherwise altering the available software, hardware, and languages. In my observation, this discouraged really innovative usage of computing resources. Time-sharing helped, but unpredictable down times and noisy phone lines still discouraged regular creative use. I believe the great successes of micros, minis, and workstations are due primarily to the fact that users are freed from the dictatorship of central processing administrators. My principle has been that *you don't use what you don't own.* By this I mean that there is too much intellectual overhead in constantly relearning updated systems to encourage creative computing. This is clearly the benefit of systems scaled to departments or individuals. It is also a principle whose recognition I would strongly encourage as an adjunct to the Eddy report.

The issue of supercomputing is a corollary issue to this principle. Computing has essentially bifurcated into personal (mini, micro, workstation) and supercomputing. For this reason as well as the above principle, I believe the University Computer Center is becoming obsolete. Unfortunately, today's supercomputers are run by the same type of bureaucracy that runs University Computer Centers. Nonetheless, supercomputing centers not only are receiving considerable resources from NSF and other agencies, but are probably the major unexplored direction of computing which offers a significant new impact on statistical science. I strongly encourage my colleagues to explore this avenue. Currently, most of these resources are being used by physicists and fluid dynamicists. If this usage pattern persists, it is likely to become institutionalized, making access by statisticians more difficult in the future. Thus the time is at hand to become involved.

The comments by Eddy and his committee on support staff to install, maintain, and develop the computing resources is clearly one that needs to be re-echoed. The granting agencies have not intended the NSF SCREMS and DoD URIP to provide for maintenance and support personnel. These are key elements of a computing activity and clearly must be provided for by either department or university resources or by other contract support.

The final thought I'd like to mention relates to the Committee's recommendation to continue SCREMS and URIP. The DoD University Research Instrumentation Program was a 5-year program begun in fiscal 1983 and scheduled to end in fiscal 1987. Decisions for fiscal 1987 are now already made. Thus as presently planned, DoD URIP will not be continued. It was supplemented in fiscal 1986 with the DoD University Research Initiative (URI). URI contains a

provision for instrumentation, but provision of instrumentation is in competition with many other objectives of URI. It seems to me that the ability to acquire significant instrumentation resources from the DoD is now substantially diminished and likely not to return unless political pressure is used.

In summary, my points are:

- You don't use what you don't own.
- Standardization aids communication and algorithm exchange and thus is highly desirable.
- Movement from minis to workstations seems prevalent and also desirable.
- The time is now to explore supercomputing.
- Provisions for maintenance and support personnel are key elements of planning.
- Acquisition of equipment (other than supercomputer access) from federal sources is likely to be somewhat more difficult in the future.

I mention in closing that Wegman (1986) contains some personal perspectives on how computing relates to statistics.

ADDITIONAL REFERENCE

WEGMAN, E. J. (1986). Midcourse musings. *IMS Bull.* **15** 238–241.

Rejoinder

William F. Eddy

I would like to thank all of the discussants for their uniformly positive comments; I wish they had been members of the Workshop. I would also like to take this opportunity to publicly thank all of the members of the Workshop for their hard work that led to the report. I am sorry that the publication schedule prevented them from having the opportunity to join me in this response.

The activities begun in the Workshop are continuing; we are organizing a session at Computer Science and Statistics: 19th Symposium on the Interface to be held at Temple University on March 8–11, 1987. This session will provide what I hope becomes a continuing public forum for discussion of both technical issues and some of the broader matters raised by this report.

SUPERCOMPUTERS

Prem Goel, David Scott, and Ed Wegman all mentioned the National Science Foundation (NSF) (and other) supercomputer centers. I agree completely with Wegman's principle: *you don't use what you don't own.* I also agree with Goel's recommendation to let others provide access to these centers. And I agree with Scott that supercomputers are not a panacea. Personally, I am not sure that supercomputers will have much positive impact on statistical research. Supercomputers are very good at doing linear algebraic calculations but are distinctly not cost effective for many other kinds of calculations.

An hour on my local Cray X-MP/48 is currently valued at $1000. An hour on a Cray X-MP is roughly equal to 30 hours on a VAX 11/780 if the particular problem is not amenable to vectorization; if the problem is totally vectorizable a Cray hour is roughly equal to 300 VAX hours (although, see Dongarra and Hewitt, 1986, for a report on a particular calculation where a Cray hour is roughly equal to 7000 VAX hours). I am able to buy a VAX 11/780 equivalent for $6000. I would much rather have 10 such machines than 60 hours on a Cray because they will support a much wider range of computing activities (and they will last longer). The most common statistical use of supercomputers is for large simulation experiments. It is interesting to note that if such an experiment is not vectorizable but is decomposable into several parallel computing tasks (such as one for each independent sample), the ten VAX equivalents together operate at roughly one-third the speed of the Cray.

The major negative impact of supercomputer centers on statistics comes from the developing sense within the funding agencies that general computational needs are being satisfied by the national centers together with a few dollars in individual research grants to buy inexpensive workstations. While workstations and supercomputers can satisfy a large fraction of the needs, there will continue to be highly diverse and specialized needs for other sorts of computing equipment. Section 5.2 of the report tried to point out that some aspects of research in computational statistics are more like computer science than like applied mathematics; I believe that graphics and parallel computation both provide fertile ground for statistical research and both require specialized equipment that is expensive and difficult to acquire without substantial external support.

STANDARDIZATION

Doug Bates, Andreas Buja, Ed Fowlkes, Jon Kettenring, David Scott, and Edward Wegman have all referred to our recommendation concerning