Bulletin of the Seismological Society of America, Vol. 80, No. 4, pp. 1046–1047, August 1990

REPLACE THE HYPO71 FORMAT?

By W. H. K. Lee

Havskov (1990) presented several reasons why the HYPO71 format should be replaced and proposed the Nordic format for consideration. I agree wholeheartedly, because I happen to be the person who first questioned the HYPO71 format. Readers may wonder why we make a fuss over a trivial matter like a data format, when it is a simple matter nowadays to reformat any data or modify any program to read any data format. One may argue that it is the data content and its quality that really matters and not its format.

Computing is vastly different than it was 20 years ago. Looking back, HYPO71 is a crudely written computer program, and its data format appears to be unnecessarily rigid for humans. This is not surprising because the HYPO71 format was designed for punched card machines! I did not invent the HYPO71 format; it was borrowed from Jerry Eaton's HYPOLAYR (Eaton, 1969). My sole contribution was to append the signal duration at the far end of the card. Most of Havskov's objections to the HYPO71 format went through my head. However, a look at a roomful of cards that were already in Jerry's format quickly changed my mind. Besides, it is not easy to improve over Jerry's format for the punched card machines.

HYPO71 was written for a dense, telemetered network of high-gain seismometers. A "cookbook" style manual was written so that my part-time student assistants could follow the "recipes." HYPO71 was released as a USGS Open-File Report (Lee and Lahr, 1972, 1975) to let taxpayers know that their government bureaucrats had done some work. As is common practice in science, anyone who asked for one got a copy. About 1000 copies were distributed this way to the great annoyance of my bosses because I kept asking for money to make copies.

I never dreamed of the comments I would receive over the years. Here are two samples of users' feedback:

> "I just made a small change in the code, and your @#$%!&* HYPO71 crashed!"
>
> "Your damn HYPO71 is useless and can't even locate an earthquake out in the ocean for my three-station network in Timbuktu."

On the other hand, a few die-hard fans insisted that I should not tamper with HYPO71: "If it isn't broken, don't fix it".

Given its many limitations as documented by Havskov, I am surprised that anyone would still want to use HYPO71. I suspect that most people don't. HYPO71 had been modified by so many seismologists and programmers that I have on several occasions asked them not to credit me. Furthermore, many new earthquake location programs have appeared since HYPO71. As to its "cursed" format, it somehow became a *de facto* standard. Fifteen years ago, Al Lindh came to my office and asked me what I thought was my "greatest" contribution to seismology. I obviously disappointed him with my answer. He then showed me the punched cards which he just received from another institution. He said it was amazing how the holes line up nicely with our cards, and this was my "greatest" contribution. I think Al was at least partially correct: everyone would agree on the "greatest" part, but many would substitute "damage" for "contribution".

Because research needs and tools change with time, we will always need new data formats. As Havskov pointed out, the Nordic format has been a success for meeting

the needs of the Nordic groups. It will no doubt please other people that have similar needs. But for those whose needs are different, the Nordic format could be a little painful too. For example, using a code in Column 80 to show what kind of data are to follow next is great for the programmers who write computer code, but can be inconvenient and prone to error for those who enter the data by typing. A numeric code is also not obvious as to what kind of data is being dealt with. Better methods may be either to use a key word at the beginning of a line as practiced by several seismologists now, or to define data structures as advocated by Ward (1989).

Experience has shown that it is not yet possible to design a data format that will be useful in all situations. If it is too general, then it will be clumsy; if it is too specific, then it will be restrictive. If the design of a data format is left to a committee, then the result most likely will be a format that everyone hates. I am grateful to Havskov for publicizing the problem, but data formats will keep evolving in this changing world. Every now and then a particular data format becomes a standard, lasts a while, and is swept away. And I don't think the fate of the HYPO71 format will be any different.

References

Eaton, J. P. (1969). *HYPOLAYR, a computer program for determining hypocenters of local earthquakes in an earth consisting of uniform flat layers over a half space, U. S. Geol. Surv. Open-File Rept.*, 155 pp.

Havskov, J. (1990). Replace the HYPO71 format, *Bull. Seism. Soc. Am.* **80,** 1043–1045.

Lee, W. H. K. and J. C. Lahr (1972). *HYPO71: a computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes, U. S. Geol. Surv. Open-File Rept.*, 100 pp.

Lee, W. H. K. and J. C. Lahr (1975). *HYPO71 (revised): a computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes, U. S. Geol. Surv. Open-File Rept. 75–311*, 116 pp.

Ward, P. L. (1989). *SUDS: Seismic Unified Data System, U. S. Geol. Surv. Open-File Rept. 89–188*, 123 pp.

United States Geological Survey
345 Middlefield Road
Menlo Park, California 94025

Manuscript received 22 March 1990