UNIVERSITY OF WASHINGTON

DEPARTMENT OF OCEANOGRAPHY

Seattle 5, Washington

Technical Report No. 52

CONFERENCE ON RAPID DATA-HANDLING IN OCEANOGRAPHY

Held at Lake Wilderness Lodge
Maple Valley, Washington
14 - 16 November 1956

by

Robert G. Paquette

Office of Naval Research
Contract Nonr-477(01)
Project NR 083-138

Reference 57-11
May 1957

RICHARD H. FLEMING
Executive Officer

ABSTRACT

The Conference on Rapid Data-Handling in Oceanography met near Seattle, Washington, on 14-16 November 1956 to determine if high-speed data-processing methods could reduce the backlog of unreported and unanalyzed data and ease the burden of data-processing for institutions collecting routine oceanographic data.

It was determined that such methods are at least on the borderline of being desirable, feasible, and economical. It was recommended that oceanographic institutions start to use electronic data-processing for the routine data so widely collected as to be termed "classical". It was further recommended that institutions beginning to process data by high-speed methods use the type of data-handling equipment locally available to them, with the one proviso that the input-output medium of the machines be readily interconvertible with those of other machines. It was felt the institutions need not be too concerned at first with complete standardization of units, quantities reported, or machine coding. It was proposed that a committee be formed to help bring an early agreement in these practices as the methods become established.

A number of participants gave details of their experiences with electronic date-processing methods as applied both to special problems and to problems and techniques of far-reaching general applicability. Perhaps outstanding among the latter was the description of the FOSDIC system for the readily accessible high-density storage of IBM card images on 16 mm film, using a machine recently developed by the U. S. Weather Bureau. A description of the data-handling methods of the

U. S. Navy Hydrographic Office was followed by a plea from that office for oceanographic data from all sources, and a suggestion that the Hydrographic Office might serve as a central repository for oceanographic data. These latter two suggestions met with some disagreement among the participants.

It was foreseen that oceanographic data would be collected at increasing rates in the future, eventually synoptically and serving predictive purposes analogous to meteorological data. New instruments will be used which will tend toward in situ measurement and analog recording. High-speed data systems will be needed under such circumstances, and instrument engineers were cautioned to keep in mind the possibility of converting analog outputs to digital form at the instrument in order to reduce the delay and cost of converting data for storage in the digital data-storage system.

The unattended and unskilled use of instruments will require provision for adequate standardization in the field, a precaution even now not receiving sufficient emphasis.

ACKNOWLEDGMENTS

This conference was made possible by special funds provided by the U. S. Navy, Office of Naval Research as an addendum to Contract Nonr-477(01) with the University of Washington. Mr. Richard C. Vetter of ONR provided the initial impetus for the conference and suggested most of the participants. The writer also wishes to thank the following: Mrs. Betty B. Dimmitt of the University Office of Short Courses and Conferences who made the preliminary arrangements for the lodge; Messrs. Cuthbert M. Love and Donald S. Hanson who took care of many of the physical arrangements and some of the notes; Professors Richard H. Fleming and Clifford A. Barnes and Mrs. Edith L. Haselwood who read and criticized the manuscript; and last but not least, Mr. James M. Snodgrass who organized and directed the proceedings, invited some of the participants, and wrote the "Chairman's Remarks".

TABLE OF CONTENTS

INTRODUCTION

Means of increasing the speed of processing oceanographic data was the subject of a conference held at Lake Wilderness Lodge near Seattle, Washington, on 14-16 November 1956. The conference was sponsored jointly by the Department of Oceanography of the University of Washington and the U. S. Navy, Office of Naval Research. The 21 delegates were oceanographers or other persons directly concerned with oceanographic instrumentation, with the handling of data, and with the use of electronic data-processing machines. Mr. James M. Snodgrass of Scripps Institution of Oceanography acted as chairman, and Dr. Robert G. Paquette of the University of Washington served as local chairman.

Because the meetings were informal and because so much material was contributed during the discussions, the logical presentation of the proceedings cannot always be chronological, nor can credit always be given to individual delegates for many of the ideas presented. A resume of the formal talks given at the conference is included in the appendix.

SECTION 1

THE PROBLEM

The premise upon which this conference was based is that oceanographic institutions have collected, and are collecting, more data than they can process with the funds available. That the situation exists was not denied by the delegates of the oceanographic institutions represented at the conference. That the situation is not confined to this continent nor to the present time is seen by inspection of the literature describing the famous deep-sea expeditions. The time lapse between completion of a cruise and publication of the results ranges from a few years to several decades.

An outsider might well ask why the budget which provides for the making of measurements does not also provide for the efficient processing of the data. The answer lies principally in the very high cost of operating sea-going ships. The cost of getting a vessel to a chosen point in the ocean is so high, and the cost of making additional measurements and observations is so low in comparison, that oceanographers feel impelled to make as many measurements as are feasible, even though it may be practicable to work these up in a very limited fashion only. Additional impetus is given to this tendency by the probability that each oceanographic station is a unique opportunity. The oceans are vast and are changeable with time, weather, and season, so it is unlikely that position and time of year will be duplicated within 60 miles and within 30 days on any subsequent cruise in the near future.

A factor contributing to the problem is the relatively high proportion of exploratory research being done, which is due to the fact that oceanography is young as a science, perhaps analogous to the better-established sciences one hundred years ago. There is a tendency to examine the exploratory measurements for salient phenomena only and then go on to new exploration. An examination of the history of science will show that other sciences proceeded in just this way in their early stages. Justifiable though this procedure may be, it does add to the backlog of unprocessed data.

Still another factor retarding the processing of data is the necessity for experienced judgment in the evaluation of measurements. Oceanographic instruments, for the greater part, must sample and measure unseen by the human eye and under conditions so severe as to invite failure. Instruments are prone to fail in unexpected ways which may require considerable ingenuity for detection.

Data must be published if they are to attain their maximum utility. if attention is directed to publishing observed data only, with a minimum of calculated results, the task is still a formidable and a tedious one. There then still remains the problem of using the data to describe the behavior of the sea. Beyond a certain point, which is reached very quickly, a study of the variables of the sea requires treatments which are nearly impossible to handle by manual methods because of the great number of data involved.

MORE DETAILED ASPECTS OF THE PROBLEM

In considering methods of speeding the processing of oceanographic data, the material may be classified as follows:

1. Types of data which now exist in large numbers and which are being produced in large numbers
2. Types of data which are specialized to such an extent that equipment for processing them will be needed by only a few organizations or one or two central agencies
3. Data to be taken in the future, the nature of which can be predicted only in part.

1. Classical Data

In the first class are routine observations at the surface and at depth of such variables as salinity, temperature, dissolved oxygen, and inorganic phosphate, together with position, time, date, meteorological information, and sea surface conditions. Dr. John Lyman of the Hydrographic

Office stated that this office, which is recording all available data of this type on IBM cards, has about 1 million cards not including Japanese data, each card representing an individual set of observations at one depth and at one time. The Japanese data number about 450,000, and many data still remain in the files of various organizations. From the data on IBM cards, the Hydrographic Office uses an automatic machine to calculate the derived quantities: density anomaly, specific volume anomaly, dynamic height, and also sound velocity; and then stores these results on cards.

Other organizations make these calculations by manual methods, then plot various diagrams as an aid to visualization, and finally deduce, for some particular region and period in time, some information about such things as the average circulation of the oceans, the way in which properties vary, and the way in which these properties are likely to affect other things dependent on them, such as weather, the biological regime, etc.

These observations are continuing, probably at a greater rate than ever before; and these are the observations that are saturating the facilities of oceanographic institutions and for which there is urgent need of rapid processing methods.

2. Specialized Data

The second class, specialized data, can be represented by several different types of measurement. Bathythermograms, which are line traces of temperature versus depth on smoked glass slides, are obtained in large numbers by Navy ships and by research vessels and are already being processed by a central agency, the Hydrographic Office. The chief objection to this procedure at present seems to be the delay involved before photographic copies are returned to the scientific group submitting the slides.

There is a likelihood that this delay will be greatly reduced in the near future; if it is not, it was felt that individual organizations might find it necessary to set up their own copying facilities.

Another type of data being taken in considerable numbers is acoustical measurements. Because this is being done by relatively few groups and is not of such general oceanographic interest, it appears that the problems of handling are best left to the groups involved.

Other problems require an individual method of solution, and examples of the value of data-processing machines in solving such problems were presented during the discussions. These were of particular value in illustrating the capabilities of machine methods for those unfamiliar with them. For example, Mr. Richard C. Vetter described Pierson's analysis of two-dimensional power spectra of waves. The data for this problem were derived by aerial stereophotography in conjunction with observations from ships, and the analysis was a sizeable problem for a large computer. There was no suggestion, however, that consideration of the handling of such individual problems lay within the province of the conference.

3. Future Data

The prediction of the future is of particular importance in that methods now devised for storing data should be compatible with work that is likely to be done in the future. In addition, decisions on how to treat data of a certain type now might well be influenced by the volume of such data predicted for the future.

It appears likely that the classical observations will increase greatly in volume as new measuring methods are devised and as more ships and other observation platforms, such as buoys, become available. Other observations, both old and new, will increase in number. It is probable that new instrumentation will trend toward analog outputs, such as graphical presentations or frequency presentations on magnetic tape. There is a possibility that, like the Weather Bureau, oceanography will have synoptic and predictive functions which would require rapid collection of a great deal of data, and rapid and frequent production of charts of oceanographic conditions.

It seems certain that the volume and variety of oceanographic data in the future will be such that machine processing will be required. In planning practicable systems of rapid data-handling for present use, the proviso should be made that systems now devised be compatible with what is likely to be done in the future, and that the medium of data storage should have flexibility with respect to inclusion of new data and data of increased accuracy. It should be essential that we do not lose forever accuracy and detail now measured but as yet not found useful.

NUMBER OF OCEANOGRAPHIC DATA UNDER CONSIDERATION

It has been noted that the Hydrographic Office has about one million IBM cards representing classical oceanographic data from all sources other than the Japanese, and that there are enough Japanese data for 450,000 cards. The number of data not available to the Hydrographic Office is not known. It is estimated that the operation of the University of Washington, as a medium-sized oceanographic institution, produces classical data at the rate of 30,000 to 50,000 individual items per year, not counting biological measurements, bathythermograms, towed thermograph records, directly measured currents, or GEK observations. This estimate is in agreement with that made by the Pacific Oceanographic Group at Nanaimo, B. C., and would correspond roughly to 21,000 to 35,000 cards of the type used by the Hydrographic Office, assuming four cards per depth per station (Master; Detail #2; and two Detail #1, one of which is used for standard depths) and one card only for a between-station surface observation. Inclusion of the other observations mentioned might well require a fifth card or more for each station as the data is used for additional research.

NATURE OF CLASSICAL DATA AND THE TYPICAL PROCESSING OPERATION

The classical measurements at the present time are obtained mostly by discrete measuring procedures. Sampling bottles are lowered to discrete depths on a wire; samples are drawn into bottles and later analyzed chemically; temperatures are read; anemometer pulses are counted; positions are calculated; cloud cover is estimated; and these data are recorded as numbers on tabular data sheets. Some of the data are then subjected to simple calculations; for example, spectrophotometer readings are converted to concentration of phosphate, or thermometer readings are converted to corrected temperatures. Most of this work is done aboard ship except for the determination of salinity, which is now commonly done ashore.

The depths calculated from pressure-sensitive reversing thermometers exposed to the sea are examined for anomalies which may indicate personal errors or instrumental malfunctions, such as tripping of the sampling bottle at the wrong depth, perhaps while the bottles are being lowered or hoisted. This usually requires a graphical comparison of measured wire lengths and thermometric depth, and also requires some experience and judgment on the part of the scientist.

At present the other water properties are plotted against depth and examined for anomalies that may be due to personal errors or malfunctioning of instruments. It is the practice of the University of Washington, and probably other institutions, not to reject apparently anomalous observations unless they can be proved wrong. Frequently the error is found and corrected.

When all the data appear reliable, the curves are used to interpolate and the water properties at standard depth are tabulated. Interpolation of salinity and temperature is checked further by means of a plot of temperature versus salinity, which is required to show a characteristic type of continuity. Obviously the latter cannot be done aboard ship if salinities are not immediately available. On some cruises, Scripps Institution of Oceanography is now determining salinities aboard by titration; Woods Hole Oceanographic Institution is now beginning to use a special salinity bridge at sea; and the University of Washington and the Pacific Oceanographic Group expect to begin using similar bridges next summer.

As an aid to understanding, some of the tabular data are now plotted as isopleths in horizontal or vertical sections of the sea.

With the above qualification regarding salinities, it would seem possible to have the treatment of the data proceed to tabulation of properties at standard depths aboard ship, since there is usually considerable time available between the stations at which observations are made. This is contingent on having enough skilled personnel aboard; conceivably, it might be possible that the data could be punched on IBM cards aboard.

Assuming that the data reach shore analyzed to this degree (although this is an ideal seldom attained in practice up to now), we find that there are a number of requirements which should be satisfied:

1. Certain other groups will have an immediate need for the data, in whatever form they are, but the more nearly complete the better. In the case of the joint North Pacific Survey, NORPAC, such groups were those whose surveys overlapped or adjoined.
2. Temperature-salinity-depth data must be converted to densities and "dynamic heights". At the present time this is done manually except by the Hydrographic Office, which is prepared to take corrected data at observed depths and carry through all the remaining calculations to dynamic heights by machine.
3. The data should be plotted on various diagrams for purposes of interpretation.
4. Routine data reports must be prepared for sponsoring agencies, libraries, and others who require access to the data but without great urgency. More complete reports in a different form must then be prepared for printed data records such as "Oceanic Observations of the Pacific." The latter serve the purpose of archival storage. All such reporting is now being done by typing and duplicating.
5. All of this should be completed in a reasonable time.
6. The organization should be prepared to meet unexpected requests for part or all of the data at any time, which in some cases may require a tedious hand-transcribing job.

In practice, the attainment of these objectives has been imperfect because of lack of time. Often the necessary diagrams are not all drawn, and interpretation of them has been sketchy. In the case of Operation NORPAC, it was only by an extraordinary effort that Requirement 1 was met in about 2 months after the end of the cruise, and Requirement 4 in about 6 months.

At this point it is usually time to start intensive planning for the next summer's heavy cruise schedule. Most of the effort has been expended in handling data and very little interpretation has been done. Any further investigations which require manipulating data or statistical treatment are rendered almost impossible due to the quantities of data to be handled. It would appear that machine-processing could relieve skilled personnel from routine duties for interpretation and analysis functions, and could also simplify analysis and permit investigations not now possible.

The practicability of using high speed machines to solve the problem of processing oceanographic data will be treated in the next section.

SECTION 2

PRACTICAL QUESTIONS REGARDING ELECTRONIC DATA-PROCESSING

This section is organized in question and answer form since this seems the most useful style to summarize the thoughts of the conference on the question of the practicability of the use of machine methods. It is impossible to do full justice to the contributions of individual delegates because so much material was presented during the informal discussions and unfortunately the notes of the proceedings fail to include any but those given formal opportunity for talks. The questions outline the main topics discussed and, in general, the answers represent the consensus of opinion of the conference.

1. ARE THERE ENOUGH DATA TO JUSTIFY MACHINE PROCESSING?

The answer was "yes". This is especially true if the stencils or master sheets for final data reports are prepared by machine also, so that no errors are introduced at any point by human handling. There is no question that tabular data could be entered automatically by machine or preprinted master sheets, but verbal statements might have to be introduced by manual operation of a machine.

2. HOW SMALL ARE THE BATCHES IN WHICH THE DATA COULD BE PROCESSED ECONOMICALLY? CRUISE BY CRUISE?

This question was not answered. The writer believes it is considerably less than one year's work, perhaps as little as the data from a short cruise.

3. ARE THERE ENOUGH DATA FOR STATISTICAL TREATMENTS?

The number of data in a given month in a given one-degree square is very small, except for certain areas. For problems requiring these discriminations in time and space, there are sufficient data for only very limited types of treatments and only in parts of the oceans.

4. ARE ELECTRONIC DATA-PROCESSING MACHINES CONVENIENTLY AVAILABLE?

In general, it appears that they are. Most institutions have them close at hand, but all machines are not of the same type.

5. WHAT TYPES OF COMPUTER ARE SUITABLE FOR THE HANDLING OF CLASSICAL DATA?

Almost any type of machine can be used, from relatively slow card-operated machines with little storage like the IBM 604; to the medium-speed machines with considerable storage like the card-operated IBM 650; the Datatron, which can accept cards or tapes; and a number of others. The very large and fast machines are neither necessary nor economical at present.

6. ARE THE INPUT AND OUTPUT MEDIA OF THE VARIOUS TYPES OF MACHINES READILY INTERCHANGEABLE?

Unfortunately they are not. To transform the magnetic or punched-paper tape of one machine into that used by another requires a special machine which usually does not exist. Such a machine may be built, but it is likely to cost something on the order of $20,000. Most manufacturers of computers have made machines for transforming their input-output media to punched cards, and most transformations between machines can be made via the punched IBM card.

7. THEN THE EXCHANGE OF DATA BETWEEN ORGANIZATIONS USING ELECTRONIC DATA PROCESSING USUALLY WOULD HAVE TO TAKE PLACE BY THE MEDIUM OF PUNCHED IBM CARDS?

Yes, this would be true in general. It would be necessary for an organization using a machine with another input-output medium to have access to a machine for interconverting with cards.

8. MAY DATA BE REPRODUCED MECHANICALLY FROM THE MACHINE OUTPUT MEDIUM IN OTHER FORM THAN CARDS OR TAPE?

For special purposes there are various methods of portrayal possible:

a. Line printers, which type an entire line of characters simultaneously at about 150 lines per minute. The results may be duplicated by photo-offset or multilith, or mimeograph masters can be prepared directly.

b. Ordinary electric typewriters, which type one character at a time. These are slow compared to the line printer. Again, the record may be reproduced by the usual processes.

c. Interpreters, which type at the top of the punched card, an alphabetic and numerical interpretation of the punches.

d. Tape printers.

9. WHAT ADVANTAGES AND DISADVANTAGES ARE POSSESSED BY THE VARIOUS COMPUTER INPUT-OUTPUT MEDIA FOR ARCHIVAL STORAGE OF DATA AND FOR USE BY A COMPUTER?

The advantages of tape over cards lies in the greatly increased speed of input and output in the computer; in the reduced bulk of storage; and, in the case of magnetic tape, the longer storage life. Some paper-tape inputs are very slow, notably the ALWAC. The speed and storage space are estimated roughly as follows:

	Speed	Storage Space
	decimal digits/sec.	cu. ft/1000 cards equiv.
Magnetic tape	600 - 5,000	0.0007
Paper tape	70 - 900	0.01
IBM cards	330 max.	0.35
FOSDIC	15,000	0.002

The storage life of ordinary IBM cards was estimated at about 15 - 30 years. Special card stock could be used both for longer life and resistance to humidity (Stern, Harmon). Paper tape would present similar problems. Both magnetic and paper tape have the advantage that available storage space for a block of data is not limited as it is on cards. When more than one card is used for storage, a considerable portion of each card must be wasted in repeating identifying information. Cards have the advantage of being readily interpreted visually; with paper tape this is difficult and with magnetic tape it is impossible. Magnetic tape has a serious disadvantage in that it may easily be erased accidentally, and with much data to be handled this is certain to occur occasionally.

The FOSDIC system has been developed by the U. S. Weather Bureau as an archival storage system for punched cards. Cards are reproduced on 16 mm film at a density of 11 cards/inch. A rapid-searching system will scan the film and transform chosen frames to punched cards again for actual computation. This possesses advantages of small storage space and long life, and removes the chief objections to punched cards as a working medium.
A machine which will use the FOSDIC record directly for input to a computer is now envisioned.

10. MAY GRAPHS AND DIAGRAMS AND OTHER AIDS TO VISUALIZATION BE DRAWN BY MACHINE DIRECTLY FROM CARD OR TAPE DATA?

Machines have been made to plot points, but the few which draw curves require extremely close spacing of points. The space density of oceanographic information is so low that it is difficult to conceive of a machine having the necessary interpretive functions to interpolate a curve such as an isopleth in an array of points.

Mr. Walter Bohan of Cook Research Laboratories presented an example of an automatically-produced digital system of data storage which has a visual element inherent in it. This example presented the twice-daily soundings by radiosonde every hundred feet of height up to about 40,000 feet, for one month with an accuracy of 0.25° C. and the fluctuating level of certain isotherms could readily be seen. The record was approximately 8 by 16 inches. This system is best adapted to presenting data collected at uniform time or space intervals. No machine exists at present for getting the data back off the record automatically for electronic data-processing.

11. HOW MUCH STANDARDIZATION OF UNITS, QUANTITIES REPORTED, AND ARRANGEMENT ON THE STORAGE MEDIUM, IS NECESSARY TO MAKE IT POSSIBLE TO EXCHANGE DATA AMONG ORGANIZATIONS?

Standardization is very advantageous but it will be obtained slowly and with difficulty. If organizations were to wait for extensive standardization they might not start automatic processing for several years. There is already considerable standardization in quantities reported and in units for classical oceanographic data. As to the arrangement of information on the storage medium, the cards used by the Hydrographic Office are the only existing example. This may not be the best system, but it will serve as a basis for starting. Better arrangements can be devised later. Transformation from one arrangement to another can be accomplished quickly by machine. Transformation of data into different units may be accomplished automatically as the data are being reproduced, and an organization may be furnished a set of data in card or tape form with units and arrangement as they wish.

It is desirable that as much standardization as possible be achieved before the storage of data on cards or other media becomes too large. It was suggested that a committee be formed to make recommendations on standardization of units and arrangement and on the compatibility of machine input-output media. The committee should also exchange information on available methods of processing data and serve in a general advisory capacity to oceanographic institutions using electronic data processing.

12. CAN THE DATA STORAGE SYSTEM BE VERSATILE SO AS TO ACCEPT ADDITIONS AND MODIFICATIONS SUCH AS NEW TYPES OF DATA, OR OLD TYPES OF DATA WITH MORE DECIMAL PLACES?

Yes, one need only add cards as the storage on one card is used up. On tape, the data are recorded in blocks corresponding to an oceanographic station. Space may be left at the end of each block for insertion of data, and there is no practical limitation to the length of block, so that space may be provided for expansion. Tape has a great advantage over cards in that all the data are in one place and the information identifying the station is recorded only once.

On cards, each card after the first must repeat certain identifying information and, to some extent, any other data that may need to be associated with the data on that card. The Hydrographic Office, for instance, carries through 29 columns of identifying and auxiliary information from card to card.

If computers with considerable memory capacity, such as the IBM 650, are used it is not essential to re-record the auxiliary information, as the cards may be arranged in sequence and correlations between cards may be made readily. Some rather complicated questions of machine economics enter into this problem, and this is a place where oceanographers will need expert advice.

13. CAN THE ELEMENT OF HUMAN JUDGMENT, WHICH NOW ENTERS INTO EVALUATION OF THE PRIMARY DATA, BE ELIMINATED?

This question is important because upon the answer depends the feasibility of carrying through the processing of data automatically from a very early stage after collection. It is believed that the majority of participants would answer "No" to this question. It can be seen that if results obtained from the data are only slightly influenced by a single datum, and if the frequency and magnitude of errors are small, one could theoretically eliminate the human factor and tolerate the errors. Thus the decision depends upon the requirements for accuracy, but it is believed that the data are too few and the errors are too frequent and too large to permit this in oceanography today.

Probably the human element could be eliminated if frequent standardizing procedures for all the instruments and methods could be set up in the data-collection process, and if the data were subjected to an automatic quality analysis using these standards and "known" properties of the sea as criteria.

The U. S. Weather Bureau does not screen its data before punching on cards but does subject it to a quality analysis afterward based upon anomalies from the expected behavior of the atmosphere. The resulting anomalous data are not automatically excluded, but are separated for manual checking and investigation.

The Hydrographic Office starts with observed data which has already been subjected to some human evaluation but which may contain errors. It then uses machines to interpolate to standard depths and carry out all the other desired calculations, tolerating any errors which are contained in the data.

14. WHAT SHOULD BE DONE ABOUT THE STORAGE OF CONTINUOUS ANALOG DATA?

This question refers to the fact that estimates of future activities in oceanography include more widespread collection of data by in situ instruments, which produce an analog record such as a line trace of property versus depth or such as a frequency-modulated magnetic-tape record. There was general agreement that the micro-structure of the original record should not be lost by an arbitrary conversion to digital form. The gross structure can be taken off in digital form if desired but a means of storing the original detail is required. With line traces this is readily done on punched cards containing a microfilm insert which does, however, reduce storage space on the card. With magnetic tape, an edited or condensed duplicate might be stored directly.

15. ARE THERE SPECIAL PROBLEMS IN MAKING ANALOG INSTRUMENT OUTPUTS DIRECTLY ACCEPTABLE TO DIGITAL DATA-STORAGE SYSTEMS? WHAT CAN BE DONE ABOUT THIS?

Yes, there is so much variety in analog records that analog-to-digital conversions tend each to require a special expensive machine. It is much simpler to design the original instrument to produce a suitable digital output simultaneously with the analog record. This possibility should be kept in mind by those designing instruments which will accumulate much data requiring digital storage and high-speed processing.

16. WILL ELECTRONIC DATA-PROCESSING ACTUALLY REDUCE PERSONNEL REQUIREMENTS SO MORE TIME MAY BE DEVOTED TO INTERPRETATION?

Mr. Stern pointed out that historically the answer seems to have been "No" because of the tendency, legitimate or not, to put the released energy into collecting more data. This is legitimate if progress has been limited by lack of data. In oceanography the paucity of data is an important factor, but by no means has everything possible been done with existing data. Therefore, it would seem important to devote any savings in manpower to interpretive functions on existing data and enter rather cautiously into any more intensive data-collection programs.

For an organization of the size of the Department of Oceanography at the University of Washington, it is not certain that any manpower will be saved, especially during the first year of electronic data-processing.

After the first year it seems likely that there would be a saving. For larger organizations there is more probability of notable savings and for smaller organizations the converse.

17. WHAT EXTRA PERSONNEL COSTS ARE INVOLVED?

Some oceanographers must become familiar with machine methods. It is estimated that it would take most of the time of one moderately skilled person for six months to become thoroughly familiar with the machines and establish an operating procedure, assuming that use is made of an existing data-storage system such as the cards used by the Hydrographic Office. It would still take a portion of his time thereafter and, because it would be most unwise to have one man the sole repository of this skill, others would have to be trained.

18. WHAT WILL BE THE COST OF ENTERING INTO AND MAINTAINING AN ELECTRONIC DATA-PROCESSING SYSTEM TO DO NO MORE THAN IS NOW DONE MANUALLY?

Let us assume data being produced at the rate of 50,000 items per year (for a medium-sized oceanographic laboratory) and that the storage and working medium will be IBM cards of the type used by the Hydrographic Office. This would require 35,000 cards per year of which three-fourths, or about 27,000, would be punched by hand, some aboard ship. This would require about one-half of one reasonably skilled person's time aboard ship. Counting supervision of the program ashore, this might amount to one-half of a man-year. Neither the primary tabulating operations by chemists and other scientists, nor the time of those making the initial graphs and deriving interpolated data, would be eliminated.

However, there would be very notable savings thereafter in preparing reports. One to two man-years of effort should be saved in the data analysis and the secretarial groups. To balance this are the following roughly estimated costs:

Punch and verify 20,000 cards at $30/M	$ 600.00
Total machine cost for 10,000 cards at $15/M . .	150.00
Duplicate, sort, and tabulate 80,000 cards at $2.70/M	216.00
Key punch at $10/mo.	120.00
Supplies	200.00
Storage cabinets, 35,000 cards at $3/M	105.00
	$1,391.00

Labor, but not supervision, is included in these costs. In addition, there will be the costs of learning to use the machinery and some unsuspected costs of auxiliary operations. These are hard to estimate but it seems likely that the total cost will be no greater than at present, and probably a little less. Manpower costs would probably decrease somewhat after the first year.

The data would then be in a form amenable to handling by machine for interpretive purposes. There is the further likelihood that copy can be produced cheaply for direct photo-offset printing by "Oceanic Observations of the Pacific," which would be of considerable value to the group processing this data. Likewise, the data would be in a form directly acceptable to the data-storage system of the Hydrographic Office, a point which would be of value to that organization and which might lead to reciprocal services from the Hydrographic Office.

To summarize, the writer's opinion is that the University of Washington could start machine processing now, performing the same operations on the data as at present at little or no extra cost but with auxiliary advantages. After the first year, more could be done with the data at the same cost. Storage space for cards will become a serious problem within ten years but may be ignored for the present.

19. WOULD OCEANOGRAPHERS IN GENERAL HAVE TO LEARN MACHINE METHODS AND MACHINE MATHEMATICS?

If they wish to use the system to any extent, they should be familiar with the fundamentals of machine operation. This can be accomplished by attending brief courses frequently offered free of charge by the companies manufacturing the computers. In the case of persons using machine methods for research on the data, there is probably no substitute for learning both methods and mathematics, although much technical assistance is available at little or no charge from the companies. For projects of some magnitude or difficulty, expert assistance is advisable.

20. SHOULD THERE BE CENTRAL AGENCIES SET UP TO SERVE THE PURPOSES OF ARCHIVAL STORAGE? SHOULD THE SAME OR OTHER AGENCIES BE EQUIPPED WITH ELECTRONIC DATA-PROCESSING SYSTEMS TO CARRY OUT OPERATIONS WHICH ARE NOT CONVENIENT FOR INDIVIDUAL LABORATORIES?

The general feeling at the conference was that the system has not progressed far enough to recommend either central repositories or central computing agencies at this time. For at least a few years the ultimate repositories of data should be the existing printed publications: "Oceanic Observations of the Pacific" and the "Bulletin Hydrographique du Conseil International pour l'Exploration de la Mer".

Mr. Lyman pointed out that the U. S. Navy Hydrographic Office needs oceanographic data and that the Hydrographic Office is a central storage agency for such data. It was resolved that all institutions be urged to transmit copies of their unpublished routine data to the Hydrographic Office, and that the Navy be urged to provide means for making these data readily available. In return, the Hydrographic Office is willing to carry out the computations of sigma-t and dynamic height from the data, and perhaps make other data available if desired.

This resolution did not receive universal acceptance at the conference and some subsequent correspondence has raised further objections which require some explanation. The Hydrographic Office is a military agency, is responsible to the military, and has no mandate to serve library functions. There is no assurance that, in a changing military situation, all stored data might not become unavailable due to security classification or a shift of emphasis and funds to some other endeavor. Hence, it may be dangerous to rely on this agency as a sole repository.

At the present time the Hydrographic Office, as a matter of expediency, is performing services for organizations which furnish them with data. This situation could change quickly as a result of administrative decision. As little as two years ago the services rendered were very severely curtailed. Mr. Lyman points out that the obligations assumed by the Hydrographic Office in accepting data are not clearly defined, and some organizations expect too much in return. If a system of definite credits and payments could be set up, services might be rendered more expeditiously, but this would require administrative action in the Navy. High level administrative, or perhaps legislative, action might be required to give adequate assurance that the Hydrographic Office will continue to serve library functions (i.e. be a source as well as a repository for data).

There was some objection, expressed principally by Dr. John Hersey and Dr. Alan Berman, to sending unpublished, unevaluated data to the Hydrographic Office. The author believes that most oceanographers would hold with these scientists if "unevaluated" means "possibly containing considerable gross error". It is believed Dr. Hersey regards "evaluated" as meaning "ready for publication in a scientific journal like the 'Bulletin Hydrographique'". It is not known if he would be willing to compromise if the need of the Hydrographic Office were great enough for that office to assume responsibility for errors. Another objection made to the principle of sending data to the Hydrographic Office is that this may jeopardize publication in scientific journals. This objection is not evaluated.

21. WHERE AND HOW WILL BIOLOGICAL, GEOLOGICAL, AND ACOUSTICAL INFORMATION FIT INTO THE STORAGE SYSTEM?

It is emphasized that data from these fields have been neglected, not because of lack of importance but because methods of treatment are poorly worked out and methods of presentation of data have not been standardized. Probably the best way to handle these data at present would be to enter on the card or other storage medium a notation concerning their existence.

SECTION 3

CONCLUSIONS

On the last day of the conference a committee was formed to draft a set of conclusions. This committee was forced to finish before it felt that its conclusions were complete. However, in review, the writer feels that it expressed most of the major conclusions, and these are presented below with only minor changes in arrangement except for the statements with respect to the Hydrographic Office as a central processing agency. A few other conclusions are added at a point to be noted.

1. Present methods of handling oceanographic data are not satisfactory in that most of the available effort of the laboratories is expended in collection and routine processing, and little time remains for interpretation and dissemination of the information.

2. It appears to the group that the immediate problem lies with the discrete types of classical oceanographic data which are now being taken routinely at oceanographic stations and at some intermediate points.

3. We think that steps can be taken, with devices now available, to alleviate this condition. We strongly recommend that the individual institutions take these steps as soon as possible, keeping in mind the following recommendations:

   a. That the institutions use equipment which is locally available to them.

   b. If a choice exists, that they use systems whose input and output media are either directly acceptable by other types of systems or readily convertible to a common medium which is acceptable to most systems. This common medium is the IBM punched card.

c. It is economical to start one of these systems on a fairly small scale, at least to perform the functions of exchange and storage on data that is essentially digital.

d. Once this data is on cards or other media, it is amenable to certain routine computations, such as sigma-t and dynamic heights, and many interpretive operations which require handling large amounts of data or need repetitive computations.

e. We do not visualize the elimination of the conventional data report at the present time.

f. We are specifically excluding from consideration at this time the problems of handling analog data, such as continuous soundings, bathythermograms, acoustic records, wave records, etc., which we regard as special problems. These are being handled by groups and individuals with these particular interests. The use of analog data at the present time is hampered by the difficulty of analog-to-digital conversion. If this were to become easy or if the original record could be taken in a form which could easily be digitized or if it could be digitized at the instrument, these data would become much more usable.

g. Individual institutions can start using data-processing systems without being too concerned about standardization of units or coding, because these can be transformed easily by machine. Standardization is highly desirable and should be achieved before the files of these institutions become too large.

h. A central storage agency for oceanographic data exists in the Hydrographic Office. All institutions are urged to transmit copies of their unpublished routine data to this office. The Navy is urged to provide means for making these data more readily available. (This recommendation is somewhat controversial, and the reader should consult the text and the comments of Dr. Hersey for modifications.) It was the feeling of the group that it is too early to make recommendations for the establishment of new central storage or central processing agencies.

i. A group of oceanographers should be formed to make recommendations on compatibility of media and standardization of units, and to disseminate information on available methods of processing data.

In addition to the items specifically agreed upon by the committee, other important conclusions reached by the conference as a whole are summarized below:

4. In the future, data is likely to be accumulated at increasing rates to satisfy a need for synoptic observations and predictions and to provide the numbers of observations needed for statistical work. New instrumentation will be developed for these purposes, and additional types of observing platforms will be used, such as buoys, towers, commercial ships, and others. Much of the instrumentation will of necessity be of types designed for unattended or unskilled operation, and the trend must be toward instruments which measure _in situ_ and record or transmit data in analog form, and often in continuous form. There will be a need for automatic processing of these data, and it would be desirable to provide for digitizing at selected points with equipment part of or auxiliary to the measuring instrument. In general

the finely detailed continuous records should not then be discarded because the fine detail is not appreciated at the moment; the continuous record should be stored for future use.

5. With unattended or unskilled operation, it will be essential to provide for sufficiently frequent calibrating operations so that human evaluation of the quality of the data can be avoided to a great extent. Even now, oceanographic methods suffer from insufficient calibration in the field.

6. A proper balance must be maintained between research on new fundamental instrumentation, which may be more or less independent of the application, and the design of instruments for application. In the latter category would be instruments adapted to fit a data system.

APPENDIX

LIST OF PARTICIPANTS

Dr. Clifford A. Barnes
Professor of Oceanography
University of Washington
Seattle 5, Washington

Dr. Alan Berman
Associate Director
Hudson Laboratories
Columbia University
Dobbs Ferry, New York

Mr. Walter A. Bohan
Meteorologist, Project Engineer
Cook Research Laboratories
Skokie, Illinois

Mr. Julius F. Bosen
Meteorological Specialist
Climatology, U. S. Weather Bureau
Washington, D. C.

Dr. Richard H. Fleming
Executive Officer
Department of Oceanography
University of Washington
Seattle 5, Washington

Dr. Nick P. Fofonoff
Pacific Oceanographic Group
c/o Pacific Biological Station
P. O. Drawer No. 6
Nanaimo, British Columbia

Mr. Alvin L. Harmon
Manager Applied Science
International Business Machines
Pacific Northwest Division
1933 5th Avenue, Seattle

Dr. John B. Hersey
Physical Oceanographer
Woods Hole Oceanographic Institution
Woods Hole, Massachusetts

CDR. D. R. Jones
Office of Chief of Naval Operations
(Op 533)
Washington 25, D. C.

Mr. Cuthbert M. Love
Research Instructor
Department of Oceanography
University of Washington
Seattle 5, Washington

Dr. John Lyman
Division of Oceanography
U. S. Navy Hydrographic Office
Washington 25, D. C.

Mr. Frank Moon
Design Engineer-Oceanographer
Seattle Development Laboratory
Minn.-Honeywell Reg. Company
Seattle, Washington

Dr. Robert G. Paquette
Research Associate Professor
Department of Oceanography
University of Washington
Seattle 5, Washington

Mr. Melvin A. Pedersen
Physicist GS-11
U. S. Navy Electronics Laboratory
Code 2233
San Diego, California

Mr. Robert O. Reid
Associate Professor
Department of Oceanography
and Meteorology
A & M College of Texas
College Station, Texas

Mr. O. E. Sette
U. S. Fish and Wildlife Service
450 B Jordan Hall
Stanford University
Stanford, California

Mr. James M. Snodgrass, Head
Special Developments Division
Scripps Institution of Oceanography
La Jolla, California

Mr. William H. ~~Stuart~~ Stewart
Electro-Data Corporation
201 Boren Avenue North
Seattle 9, Washington

Mr. Joshua Stern
Physicist
Office of Basic Instrumentation
National Bureau of Standards
Washington, D. C.

Mr. Richard C. Vetter
Office of Naval Research
Geophysics Branch
Code 416
Washington 25, D. C.

Mr. Armand J. Volta
Division of Computation
U. S. Navy Hydrographic Office
Washington 25, D. C.

SOME CONTRIBUTIONS OF INDIVIDUALS

Dr. Alan Berman

Dr. Berman discussed some of the problems of measurement and data-handling in the studies of acoustical noise and sound propagation being made by Hudson Laboratories. Many data are collected, and attempts are now being made to process these aboard ship using analog computers. Digital methods are too expensive. Up to now, data have been in the form of banks of meters, photographed every five seconds, and tabulated and analyzed manually; or in the form of continuous tape records which are cut and formed into continuous loops for frequency analysis. Dr. Berman made a plea for the routine measurement of ambient noise in the sea by oceanographers, since the numbers of such data are far from adequate. He also commented that the calculation of sound velocity from salinity and temperature was not always sufficient because factors such as air bubbles may change the velocity as much as 40 percent; therefore, he proposes that sound velocity should be measured directly with available velocimeters.

Dr. Berman also contributed ideas on the basic philosophies of data-taking and analysis, and submitted a well-organized summary of conclusions in a communication subsequent to the conference.

Mr. Walter A. Bohan

Mr. Bohan presented the elements of a unique method developed by Cook Research Laboratories for converting continuous data to a digital graphical form. In this method, only the change since the preceding measurement is recorded, and this is digitized in terms of a unit which is the "least detectable difference" in the variable. A machine exists for converting continuous records to this form and presenting the result on a sheet of paper in a readable form which is highly concentrated and has a visual aspect that, in part, takes the place of plotted diagrams. The method is best adapted to graphic measurements having a uniform time or space interval, such as repeated radiosonde soundings. The twice-daily soundings for a month are recorded on a piece of paper about 16 inches long and 8 inches wide, and the variations in level of certain isotherms are seen readily. At present, no machine exists for transforming data thus recorded into a form suitable for a computer input.

Mr. Julius F. Bosen

Mr. Bosen was most helpful at many points during the conference with his detailed knowledge of high-speed data-processing systems and their philosophy. Only a few of his contributions are listed below. He described the data-storage system of the U. S. Weather Bureau at Asheville, N. C., where data is stored on 325 million cards, and the storage is currently increasing at the rate of 30 million per year. He gave considerable information on many practical details, such as costs, storage space required, permanence of cards, and on many other points which have been incorporated into the earlier portions of this report.

He described and showed examples of the new FOSDIC system for storing data in accessible form. IBM cards are photographed on 16 mm film automatically at a density of 11 cards per inch. The resulting record may be reconverted rapidly into cards by automatic machinery which scans the film for selected information and punches out the desired cards. The system now will serve the purpose of archival storage since the medium has long life, small volume, and rapid access. All computing and analysis is to be done with punched card decks prepared temporarily for the purpose by the machine. Later, a method of making the FOSDIC record directly acceptable to the computer is envisioned.

Another topic which he discussed was the importance of not discarding the detail of continuous records by relying entirely on digital systems for storage of data. Such data have micro-features as well as macro-features. Only the macro-features should go into the digital storage system and the high-speed processing system. He also emphasized (as did Harmon later) the need for reducing the bulk of the data at the source by making instruments to produce records directly acceptable to computers while they are also making continuous records.

He noted that the Weather Bureau does not analyze data for errors before storage, but uses data-processing equipment to sort out suspicious items after they have been entered on the cards. Various checking procedures, based to a great extent on expected behaviors of the atmosphere, are used. The suspicious data are not rejected, but are investigated further.

Dr. Richard H. Fleming

Dr. Fleming gave the keynote talk and introduced the group to the problems involved, as have been outlined in the first part of this report. He described the peculiar difficulties of oceanography, some of its philosophies as a science, the nature of the data, and some of the things to be desired of data from the point of view of both analysis and dissemination. He presented the difficulties of an administrator trying to decide whether to enter into automatic processing, and he pointed out the necessity for pictorial presentations of the data as an aid to understanding and, from this point of view, the complete unintelligibility of printed data records.

Later in the conference Dr. Fleming made an estimate of future conditions in oceanography, foreseeing the collection of data at a much greater rate, probably synoptically, and using observing platforms other than special oceanographic vessels. Instruments for taking routine data would be in situ types, simple, rugged, and trouble-free, producing continuous analog records as their basic outputs. He emphasized the trend toward analog records, and suggested that oceanography in the future would have functions analogous to weather-predicting and climatology, and might issue frequent predictions of conditions useful to such groups as fishermen and mariners.

He also produced an interesting flow-diagram to illustrate the mechanisms of handling data in such a system, and also showed a diagram illustrating the various input-output media in high-speed data-processing machines and the existing routes of interconversion between media.

Dr. Nick P. Fofonoff

Dr. Fofonoff gave some details of the plans of the Pacific Oceanographic Group for rapid data-handling methods in Summer 1957. They will use the ALWAC-3E computer with Flexowriter input and output, and propose to program the entire operation from an early stage in the data-collection process through to the calculation of dynamic heights and intermediate quantities. Copy for the reproduction of data reports by photo-offset will be produced directly by the machine. The ALWAC computer at present will not accept punched cards, and the machine for interconverting tape with cards will not be available at first.

Mr. Alvin L. Harmon

Mr. Harmon gave considerable information about the capabilities of IBM machines. He too was in favor of the punched card as a preferred medium, or at least a "common denominator" between machines of different types. He noted that all of the machines in the Electronic Accounting Machines group had been used aboard ship, and some on airplanes. With respect to the exchange of data, he said that IBM had machinery for transmitting punched-card information by radio or land wire and had successfully transmitted millions of characters over long distances with few mistakes. He agreed with Volta and Bosen that analog-to-digital conversion is, in general, expensive and should be accomplished at the measuring instrument. He recommended that all of the group avail themselves of one of the one-week courses in computer principles offered free by International Business Machines.

Dr. John B. Hersey

Dr. Hersey arrived late at the conference and was not able to contribute to the earlier discussions. He mentioned some of the instrumentation with which he has been concerned, mainly in the acoustical and seismological fields. Two instruments mentioned were a high-precision echo-sounder and a new type of thermal recorder which has been engineered to print results continually at the instrument. This latter device consists of a string of thermistors, every 20 feet along a 500-foot cable, which is towed behind the ship in a more or less vertical orientation. The thermistors are scanned sequentially by a selector switch in a bridge circuit, and the temperatures are printed by an electric typewriter. Cards could be punched as easily.

Dr. Hersey emphasized the sparseness of existing data in the oceans and questioned the existence of sufficient data for statistical work. Some of his other contributions have been mentioned earlier.

In a subsequent communication he stated his opinion that the better data would be available in published form by the time it would be suitable to be turned over to a central agency such as the Hydrographic Office, which perhaps implies that that office might as well get the data from the publication as from Woods Hole. There would also be data suitable for publication, but not published because of limitations of journal space or other reasons. These data should be turned over to the central agency, he thinks, if the agency would pay the costs of transcription. Correspondingly, the agency should make a charge for furnishing data from its files.

He believes that much new instrumentation will be produced in the near future and the results may well be incompatible with some standardized data-handling system, such as that of the Hydrographic Office. Also, if instrument designers feel constrained to produce only instruments to fit into some pre-assigned data system, ingenuity may be stifled and progress in instrumentation seriously hampered.

Dr. John Lyman

Dr. Lyman outlined in some detail the way in which data are handled at the Hydrographic Office and showed examples of the IBM cards on which data are stored and calculations made. The Hydrographic Office is processing not only data taken by its own ships but all other data which it can obtain, published or not. He described a new, partly-automatized system for producing prints from bathythermograph slides, which involves first copying the slide on microfilm together with the pertinent row of data from the original BT log sheet. This microfilm copy will eventually be mounted on an IBM card together with the log sheet data and digitized data taken from the temperature-depth curve. The Hydrographic Office wants bathythermograms from the oceanographic institutions but recently has been in a poor bargaining position because of the large backlog of unprinted slides accumulated during the development of the new system.

Dr. Lyman mentioned a few of the ways in which oceanographic data are used by his division to produce results useful to the Navy. The various atlases of average ocean conditions produced by the Hydrographic Office are well known, and there are others of a classified nature. In speaking of synoptic charts, he pointed out that the Hydrographic Office already is in this field with weekly ice forecasts and sometimes other forecasts intended principally for the Navy. His office has no responsibility to fisheries and probably never would be the agency to produce synoptic charts for that group. Nevertheless, the Hydrographic Office has a great need for oceanographic data and would like to receive data from all agencies. In return, he offers to calculate and print out the usual derived quantities by machine. In reply to objections about the long time-delay involved, he said that such projects now have high priority and should be done with little delay. He pointed out that some groups expect too much in return and that if these groups were willing to pay for special services, such assistance could be furnished with much less delay.

Dr. Robert G. Paquette

Dr. Paquette acted as local chairman in making the physical arrangements for the conference. In this he was given considerable assistance by Mr. Cuthbert M. Love and Mr. Donald S. Hanson. He outlined for the group the nature of the classical data-collection process, the methods by which the data are treated, and the difficulties now being faced in making these data available to others and using it for further research. In an attempt to assess the practicability of machine processing, he made some estimates of numbers of data involved, and suggested that it appeared possible for his organization to start using machine methods in the summer of 1957.

Mr. Melvin A. Pederson

Mr. Pederson and his group at NEL have been working on the long-range propagation of underwater sound. They have been applying high-speed computers to the prediction of intensities, using ray theory. The data being used consist of 75,000 bathythermograms digitized on cards, and the results of 5,800 hydrographic stations. A computer of intermediate size, such as the Datatron or the IBM 650, has been found most useful in this work. Anderson, in the same group, is preparing an atlas of sound velocity for the North Pacific. Carsola has been measuring internal waves, using thermistors at fixed depths, and using machine methods for processing the data. NEL has developed a new electronic wave-pole and, to obtain more data on sea conditions, they are considering installing a tower off Mission Beach to continuously record temperature, waves, tides, and sound velocity. Both of these projects will produce many data which may well require machine processing.

Mr. Robert O. Reid

Mr. Reid has been using high-speed computers in the problem of predicting storm surges from tidal data. The mathematical problem is quite complicated. He has been attempting with some success to apply the method to predicting storm surges in the model of Narragansett Bay constructed by the Corps of Engineers at Vicksburg. The slowness of the method is a very serious problem in nature, since the storm surge travels across the continental shelf at about 42 knots and is even faster in deep water.

Mr. James M. Snodgrass

Mr. Snodgrass acted as chairman of the conference, organized the proceedings, and was instrumental in selecting and inviting some of the participants. Besides his many contributions during the course of the conference, he outlined some recent advances in oceanographic instrumentation at Scripps Institution of Oceanography, which are noted very briefly below.

Increasing use is being made of liquid-filled instrument cases, not pressure-protected. Such components as batteries, subminiature radio tubes, photocells, and Teledeltos paper work well under oil at very high pressures. A new type of integrating electrical recorder is an adaptation of the silver coulometer. Electrical signals may be telemetered up a single bare cable for 100 meters or so. Recently, considerable use has been made of frequency-modulated transmission using the subcarrier frequencies now standard in guided missile work, and for which commercial components are available. Several indications may be transmitted simultaneously up a single insulated steel conductor. Other components are neutrally-buoyant two-conductor cable with a breaking strength of 325 pounds, waterproof Joy connectors, and batteries activated by sea water. Transistors are not generally useful because they are sensitive both to temperature and to pressure.

Recent instruments include the following:

(1) A temperature-gradient meter for bottom sediments which is self-recording, using a servo system, and sensitive to 0.001° C.

(2) A scanning bathythermograph which automatically, or at the will of the operator, moves up and down in the water at the end of a cable towed behind the ship. Electric power is supplied from shipboard, and temperature and depth information are telemetered as frequencies on the same single-conductor insulated supporting cable. The information may then be recorded on magnetic tape for future use, or converted to voltages and recorded on an X-Y recorder. The time scale may be compressed so that a number of soundings may be shown simultaneously on the screen of a cathode-ray oscilloscope, and at the same time the frequency presentation may be played audibly. It has been found that the latter is an aid to understanding.

(3) A current meter with threshold about 0.01 knot, telemetering velocity, direction, and depth, to the surface on a single-core cable.

(4) An in situ oxygen meter based on the dropping mercury electrode. This meter has demonstrated notable fluctuations in the oxygen-depth profile in the sea.

(5) Some cheap expendable instruments are now envisioned.

Mr. Snodgrass noted that it is difficult to teach oceanographers instrumentation or the proper use of instruments, since many of them are poorly prepared in physics, and that there appears to be no thorough course offered anywhere in instrumentation suitable for oceanographers.

Mr. Joshua Stern

Mr. Stern told something of the activities of the Office of Basic Instrumentation at the Bureau of Standards, which is devising basic instrumentation for future use. One big problem is a reference system for published information on instrumentation. Recently a punched-card system, termed the PEEKABOO system, has been devised to make such information readily accessible. Punched cards are used, with 500 pilot holes which may be enlarged with a special hand punch; and recently a card with 18,000 positions has been devised. Each hole position represents a document, and each card represents a subject or a modification of a subject. The search for information on a subject having several aspects is then accomplished by stacking the cards representing the subject and its various modifiers and inspecting visually for coincident holes. Thus, a search for information on measurement of pressure in the sea might combine cards for pressure, measurement, accuracy modifier, sea water, and other modifiers to describe the method.

Among instrument developments by his office Mr. Stern mentioned a low-noise electrical cable for high-impedance work described in Bureau Reports Nos. 1388 and 4094, and an underwater velocimeter which is now being produced by the Chesapeake Instrument Corporation. The Bureau issues a number of reports which may be interesting to the group, such as one on analog-to-digital converters (No. 2755, 1953), and the semi-annual report of the Office of Basic Instrumentation.

In speaking of instrumentation in oceanography, he noted that instruments are usually developed independent of the application when real progress in instrumentation is being made. He feels that there may be a dangerous bar to progress if too much emphasis is placed on making only those instruments which will produce an output which can be fed into a data system. The same may be true if there is too much standardization in quantities measured or methods of reporting information. He was suspicious of methods of detecting errors which used the data themselves for this purpose, and pointed out that it was better to have standards for calibrating the instruments with sufficient frequency.

Mr. Richard C. Vetter

Mr. Vetter represented the Office of Naval Research at the conference. When asked what his office expected in sponsoring the conference, he replied that if it succeeded only in mutually educating the participants, ONR would be satisfied, and that we could expect no more than that the conference would reduce the number of wrong decisions made in the future about rapid data-handling.

Later in the conference he presented an outline of the work done by Pierson of New York University. Dr. Pierson unfortunately was unable to attend the conference. Pierson determined the two-dimensional power spectra of waves from data obtained from aerial stereophotographs taken near Barbados. This was a problem for a large computer. The incompatibility of input media in different machines was emphasized here, because after this vast amount of data had been programmed for one machine, that machine became unavailable and the entire job had to be done over in a different medium.

Mr. Armand J. Volta

Mr. Volta contributed helpful information and ideas at many points from apparent long experience with electronic data-processing systems. He described some of the plans of the Hydrographic Office for handling data. They plan to convert their card files to magnetic tape, are now using an intermediate-speed computer in addition to the IBM 607, and are getting a Benson-Lehner data-reader to punch cards from line traces such as bathythermograms. They are also getting an automatic plotter. For report purposes, oceanographic data are now being printed out by machine in tabular form and converted to mimeograph stencils by means of a STENOFAX machine, which produces a stencil by scanning of an electro-sensitive paper stencil material.

He voiced the opinion that satisfaction with the status quo in electronic data-processing methods could only lead to failure. Electronic data-processing systems should be considered if a problem is complex or the amount of data large. A determination of the feasibility of the electronic data-processing system for a particular problem should not be left to inexperienced persons. He favored the punched card as the common denominator between the various types of machines, especially in view of the FOSDIC system described by Bosen. He felt that future development of oceanography would parallel that of the Weather Bureau and suggested that the Hydrographic Office might possibly become the central agency corresponding to that Bureau.

CHAIRMAN'S REMARKS*

by

James M. Snodgrass

This conference on processing oceanographic data was quite successful. Most of us have come to realize that a conference of this type should have been held long ago, as many of us found that it was even later than we thought in the progress of the automatic handling and processing of data. The oceans are vast and it is inconceivable that we can prosecute successfully the biological, physical, and chemical problems without major assistance from our computer and data-processing colleagues. At the same time, we can hold out some very real comfort to those who feel that the technological advances may pass them by and leave them hopelessly isolated. Some automatic data-processing techniques which were described and illustrated at the conference demonstrate that such techniques can and do offer, to the individual who wishes to carefully contemplate and mull over every aspect of his project, a far more complete grasp of the nature of interlocking and related parameters than was beyond our wildest dreams a scant year ago. There still exists a means of working-up data on as intimate a basis as one may wish with as little or as much assistance from automation as one may wish to accept. It is apparent, however, that for an increasing number of research projects, which will inevitably require an almost astronomical number of observations and measurements, it will be necessary to accept wholeheartedly and to take advantage of every improvement

---

* Because of the physical separation and limited communication between the writer and the chairman of the conference during the preparation of this report, it was felt desirable to have Mr. Snodgrass contribute some of his personal feelings on the subject of the conference. He has written these remarks without having read this report. R. G. P.

which can be realized in the whole field of automatic data-processing and computation. Though machines are not human and human beings are not machines, it is true that the machines can and do have a remarkable ability to reject erroneous or false observations.

It has now become an established fact that, when working over a substantial quantity of data, human beings are wont to err. A relatively small amount of time spent in planning system checks in order to enable automatic equipment to reject faulty or questionable observations will pay very large dividends in the improvement of the final accuracy. Accurate results will be achieved which are impossible without the assistance of the automatic equipment.

At present, one of the most obvious and glaring deficiencies is the inadequacy of our oceanographic instruments. Whereas methods of processing and handling data have moved at rocket-like speeds, our development of instruments has proceeded at a conservative, snail-like pace. Perhaps the major difficulty is our failure to appreciate the necessity of a better integration of instruments and processing equipment. It is true that we can make substantial advances by using existing instruments supplemented by automatic processing equipment. However, this must be considered as only a stopgap measure which is grossly inadequate. Fundamentally, our instruments and sensing devices must be designed to supply output information at a high rate of speed in a suitable form to permit feeding the information directly into automatic processing equipment. In the case where data is accumulated at a relatively slow rate, the equipment itself does not have to be capable of very high speed operation, but the recorded information should be in such form that it may be subjected to direct high-speed processing of essentially automatic nature.

The individual scientist does not need to fear that he will lose all touch with the details of his project once he has committed himself to the use of automatic data-processing, because the development of the high-speed processing equipment itself makes it possible to feed back to the experimenter information in a far more useful form than he has ever had before. There is no need to feel alarm about the requirements for the form of output of an instrument, because they in nowise prevent a record from being produced in whatever conventional analog form may be desired. In fact, this is desirable in the course of many research projects, particularly during field observations, in order to determine properly future activities.

At present, the use of computers in high-speed data-processing appears to be the only hope of reducing the overly long period between the time of taking the actual observations and the publication of the data or report. Along with this very great reduction of lead time in getting out data, the high-speed methods will make possible the rapid exchange of data between scientists and institutions.

New developments, which have proceeded concurrently with the evolution of high-speed computers and data-processing equipment, appear to hold great promise from the standpoint of the general problem of archives. These new techniques make it possible to compress into an extremely small amount of space a large mass of information; at the same time, such information is readily available in a highly accessible form.