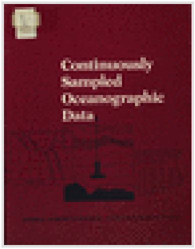


Continuously Sampled Oceanographic Data: Recommended Procedures for Acquisition, Storage, and Dissemination (1973)

Ocean Science Committee; Ocean Affairs Board; National Research Council

Pages
57

Size
8.5 x 10

ISBN
0309341833

Continuously Sampled Oceanographic Data

Recommended Procedures for Acquisition, Storage, and Dissemination

OCEAN SCIENCE COMMITTEE
OCEAN AFFAIRS BOARD
NATIONAL ACADEMY OF SCIENCES–NATIONAL RESEARCH COUNCIL

7712

Available from
Ocean Affairs Board, National Academy of Sciences
2101 Constitution Avenue, N.W.,
Washington, D.C. 20418

CONTINUOUSLY SAMPLED OCEANOGRAPHIC DATA
ad hoc Panel

Anthony Amos, *Chairperson,* Lamont-Doherty Geological Observatory of Columbia University
Arnold L. Gordon, *Chairperson,* Lamont-Doherty Geological Observatory of Columbia University
Philip R. Hadsell, National Oceanographic Data Center
Robert Ochinero, National Oceanographic Data Center

OCEAN SCIENCE COMMITTEE

John A. Knauss, *Chairperson,* University of Rhode Island
John D. Costlow, Jr., Duke University Marine Laboratory
Robert A. Duce, University of Rhode Island
Richard C. Dugdale, University of Washington
K.O. Emery, Woods Hole Oceanographic Institution
Dirk Frankenberg, University of Georgia
Edward D. Goldberg, Scripps Institution of Oceanography
John W. Kanwisher, Woods Hole Oceanographic Institution
Maurice Rattray, Jr., University of Washington
Claes H. Rooth, Rosenstiel School of Marine and Atmospheric Sciences
John G. Sclater, Massachusetts Institute of Technology
Manik Talwani, Lamont-Doherty Geological Observatory of Columbia University
Ferris Webster, Woods Hole Oceanographic Institution
Richard C. Vetter, *Executive Secretary,* National Academy of Sciences

SPONSORS

This study was carried out as part of a general series of oceanographic advisory activities sponsored by the Atomic Energy Commission, the U.S. Geological Survey, the U.S. Coast Guard, the National Oceanic and Atmospheric Administration, the National Science Foundation, the Office of Naval Research the State Department, and the Environmental Protection Agency.

CHAPTER 1
INTRODUCTION

The development and refinement of numerous types of instruments that are capable of yielding continuous or near-continuous, records of basic oceanographic parameters have challenged the oceanographer to interpret the many newly observed fine- to middle-scale features. He hopefully will understand these features and their relation to large-scale phenomena. An increased use of such instrumentation which has already produced a large quantity of data and has magnified greatly the problems of archiving and of distributing this data to secondary users is expected. The secondary user is the person who uses the data but does not collect or process it. Traditionally, he had added much information to the field of oceanography and should have these new types of data available because he lacks the resources to collect it. The availability of the continuously recorded data in an accessible form is essential. In this report a method for archiving and distributing such data is suggested. Implicit in these suggestions are recommendations to collectors of these data on the information required by the secondary user.

The principal source of this data explosion is the in situ recording electronic sensors of temperature, salinity or conductivity, and pressure (or depth) designated as the STD. A survey was conducted to determine the methods and instruments presently used in the collection of STD data. The results of this survey are given in Chapter 2.

The Ocean Science Committee, at the request of the National Oceanographic Data Center (NODC), organized a task group chaired by Drs. Amos and Gordon with the following terms of reference:

1. to describe the various techniques which are used to

obtain high resolution, digital STD output;

2. to recommend those thechiques which yield results suitable for secondary users;

3. to group, if possible, the various techniques into sets which yield comparable results and make recommendations regarding the interfiling of these data;

4. to recommend filtering and/or compression methods which should be applied by NODC after initial data reduction and quality control;

5. to recommend to NODC the types of quality control procedures which should be applied after receiving these data; and,

6. to recommend to NODC the type of information needed by users of STD data with regards to methods used to collect the data.

To meet these objectives, the task group first outlined the methods and their limitations which are presently used for collecting and processing STD data. Then, the obvious needs of the oceanographer and secondary user of these data were listed. Once this was accomplished, the task group set-out to suggest standard procedures for collecting and reporting of data which would be helpful to the secondary user. They also considered methods of archiving STD data at NODC to meet most demands. Hopefully, the procedures recommended for the collection of data will also be useful as a guide to the primary user for obtaining maximum accuracy of the STD system.

As a basic philosophy, it was assumed that little is understood about the details of the ocean, therefore "correcting" STD data should not be done in haste. The user of the data can be subjective in correcting the data, but the data filed at NODC should be corrected only by removal of obviously outrageous data.

Classical oceanographic data was chosen as the basis of at-sea calibration, assuming it was carefully collected. Classical data is reversing thermometer and deck salinometer data. Three or 4 reversing thermometers should be placed at each level sampled and should be carefully calibrated. The salinometer should be calibrated using standard sea-water. The aim is to have classical data-point accuracy in the STD data. The secondary user should judge the quality of the data on the basis of the number and depths of the classical data points, and the times of collection relative to the STD values. This would be the most meaningful quality code. Secondary users would be guided by the classical data-points and primary users should be certain that their STD is standardized by the classical data-points.

The STD not only provides information on the macrostructure of the ocean as is the series case, but it also shows the presence of a complex structure of temperature and salinity versus depth on a variety of scales. While most of these structures have limited horizontal extent, a number have very large lateral extent with very limited vertical expression. These features often separate zones or layers of differing water masses or dynamics, and so are significant to general circulation problems. The STD also shows structures on a scale approaching the distance traveled in a period close to the time constant of the sensors. Therefore, the exact character of these featrues is likely not to be portrayed accurately. These features, often called microstructure, may not be successfully studied with the standard STD. As an estimate, 1-2 meters were chosen as the smallest scale which the standard STD can study with a slow lowering rate (30 m/min). Certainly standard STD data may be used to study features of less than 1 or 2 meters vertical extent, but this should be done at the user's risk and taken from the high resolution data file (Section4). Special STD systems have recently been constructed which can observe structures on a vertical scale of centimeters. Data from such instruments would be stored in the ultra-high resolution file. The wiggles on a scale of over 5 meters are a significant part of the standard STD record and must be retained in NODC files. The many users of the STD data on a variety of scales provide a problem in that some users may want to see only equivalent Nansen data. Thus, some may want to use only a few of the STD data points and others may want to employ them all.

A single filing system would unlikely work for everyone. Hopefully the number of formats can be minimized to achieve a simple archiving system. One simple form capable of satisfying the great majority of secondary users would be included.

The measurement of T and S by *in situ* electronic sensors will eventually replace entirely the series cast and should be filed in a compatible format. At-sea standardization of the electronic sensors with classical data points will probably be necessary in the foreseeable future; especially in the salinity and pressure sensors. The deck salinometer is standardized with a known seawater sample (Copenhagen Water) before, after, and during each run of unknowns. It is unlikely that the *in situ* salinometer will be perfected beyond that point in the near future.

CHAPTER 2

SURVEY ON THE USES OF STD SYSTEMS

A questionnaire was sent to 150 users of STD equipment in November 1972. The objective of the questionnaire was to discover what systems, methods of data collection, storage, and reduction were being used throughout the oceanographic community to facilitate recommending standardized procedures to NODC. Letters were sent to various national data centers to request assistance in locating STD users in foreign countries. The response was very gratifying: 66 questionnaires were completed and additional letters were received from non-users or potential users expressing interest in this project.* There is considerable concern throughout the community with regards to the problem of STD data storage,access, and standardization.

Questions in 6 categories were asked of each user:

1. What STD system(s) are you using?

2. How do you record the raw data?

3. How do you calibrate the STD?

4. What techniques do you use in making an STD lowering?

5. How do you reduce, process, and store the STD data?

6. Do you submit STD data to NODC? And how?

Charts 1-6 summarize the data obtained from the questionnaires. A few comments about the statistical validity of the survey must be made. The specific amount of STD-type instruments

* In addition to the U.S.A., replies were received from 19 foreign countries.

that are in use or have been manufactured is unknown. Plessey Environmental Systems Inc., who make the majority of STDs in the United States, have sold approximately 450 units in the last 10 years. An estimate of the total number of STDs would be 1000 units, that would include a considerable number of estuarine, water-quality type instruments for which few responses were received, and which produce data that would not generally be submitted to NODC.

Information was received for approximately 104 systems (perhaps 10 percent of the total STD-type instruments which have been used). A large-scale user is the U.S. Coast Guard Oceanographic Unit which operates 35 systems, but they have been treated as a single system not to bias the results.* The data presented here accounts for a significant percentage of the active users of STD equipment and is a valid reflection of current techniques.

---

* The U.S. Coast Guard is an important contributor of data to NODC and we do not intend to minimize their effort, rather, the survey was conducted with the aim of determining the variations of usage techniques among a large group of different users.

GENERAL NOTES TO THE CHARTS

1. The words (number of systems) "in use" refer only to data obtained from the responders to the questionnaire.

2. All the instruments are referred to as STDs although there are many combinations of sensors in use (CHART 1C).

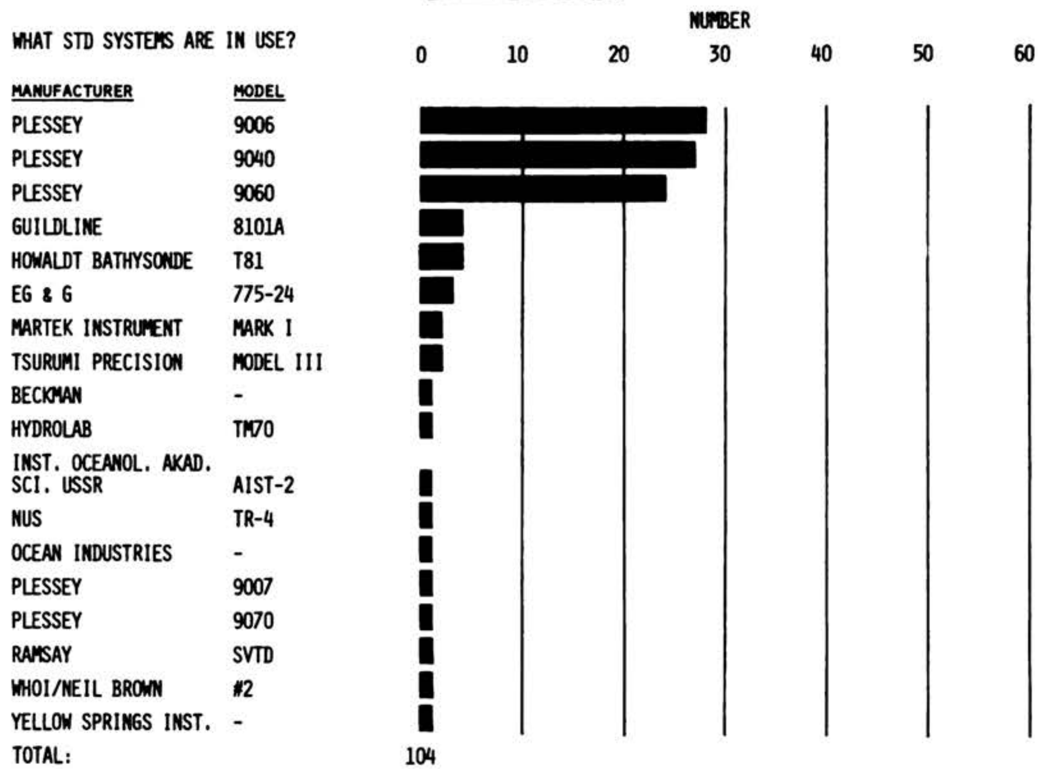
3. The numerical data may not always equal the total number of instruments in use as many categories can have repeats (i.e., recording may be made in both analog and digital form for one instrument), and not all users responded to all questions.

4. Detailed notes that pertain to the individual charts follow each section of charts.

CHART 1

INSTRUMENTATION

A WHAT STD SYSTEMS ARE IN USE?



B IS SALINITY OR CONDUCTIVITY RECORDED?

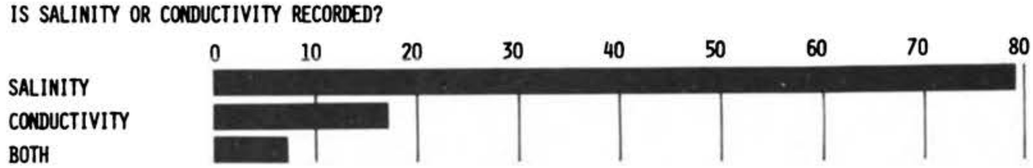
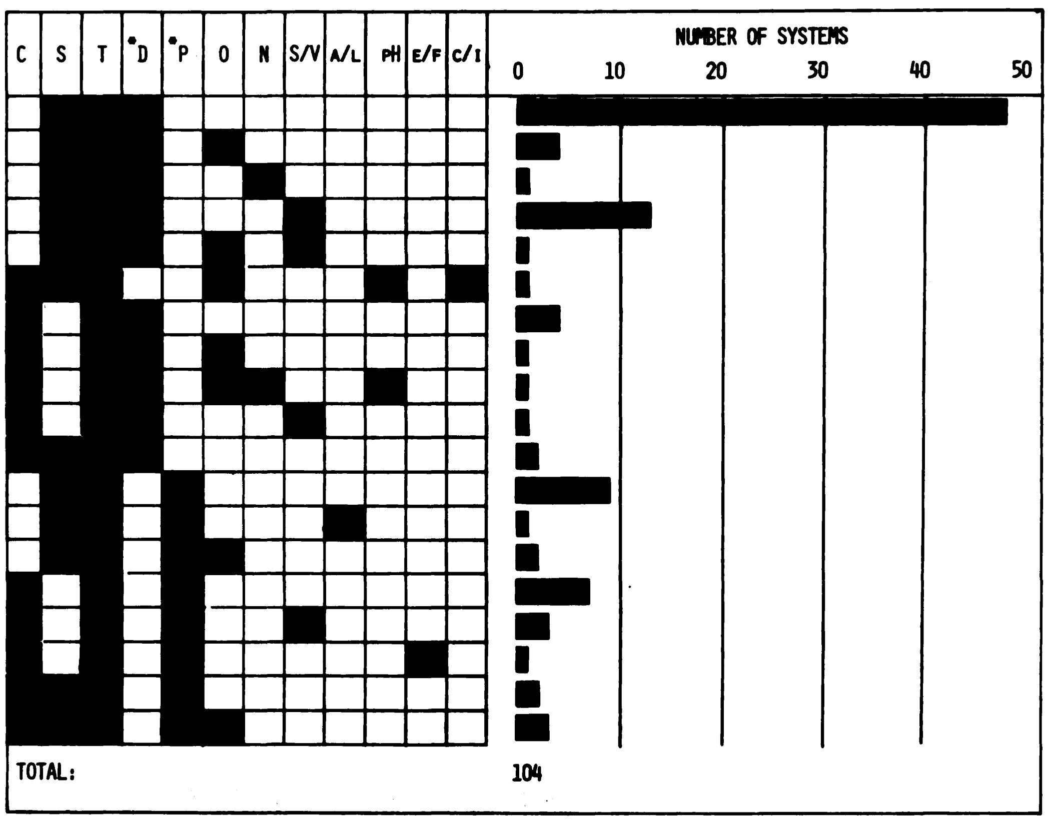


CHART 1 CONTINUED

C WHAT COMBINATIONS OF SENSORS ARE IN USE?



C = CONDUCTIVITY
S = SALINITY
T = TEMPERATURE
D = DEPTH

P = PRESSURE
O = OXYGEN
N = NEPHELOMETRY
S/V = SOUND VELOCITY

A/L = AMBIENT LIGHT
PH = HYDROGEN ION CONCENTRATION
E/F = ELECTRICAL FIELD
C/I = CHLORIDE ION CONCENTRATION

CHART 1 CONTINUED

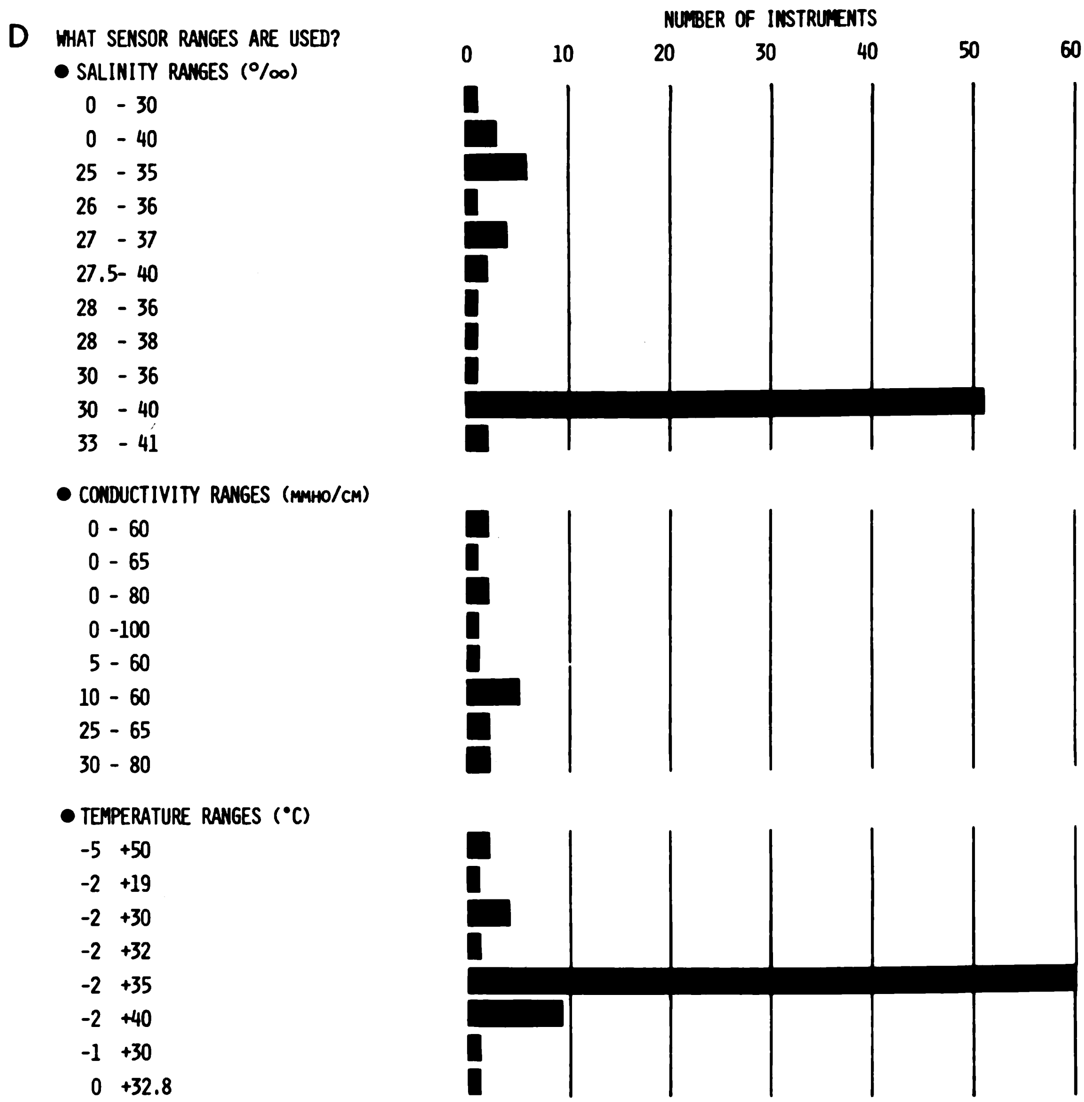
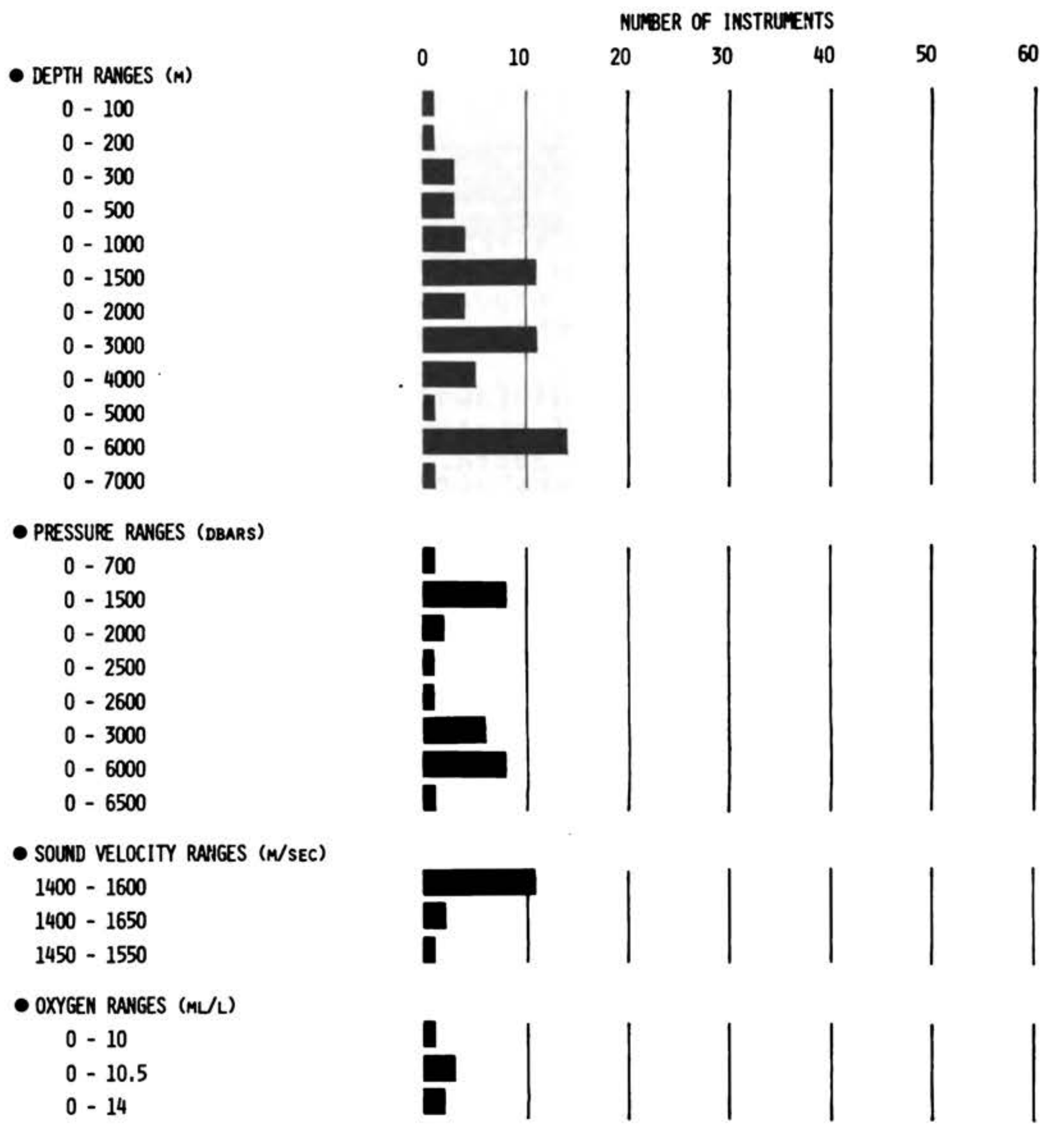


CHART 1 CONTINUED



NOTES TO CHART 1 - INSTRUMENTATION

A. *Number of instruments in use by manufacturer and model number.*

B. *All STD instruments sense conductivity. This chart indicates the breakdown of those instruments that record conductivity directly (internally or transmitted to deck equipment) and those instruments that compensate for temperature and pressure effects in situ and record salinity. The great majority record salinity.*

C. *Breakdown of what individual sensors are in use in the STD systems. All systems sense pressure as a measure of geometric depth. The chart shows how the user interprets the difference between depth and pressure (D and P).*

D. *Indicates the ranges of the sensors used. Most users have 30-40°/00 salinity sensors, -2 to +35°C temperature sensors, and 0 - 6000 m depth (pressure) sensors.*

CHART 2

RAW DATA RECORDING

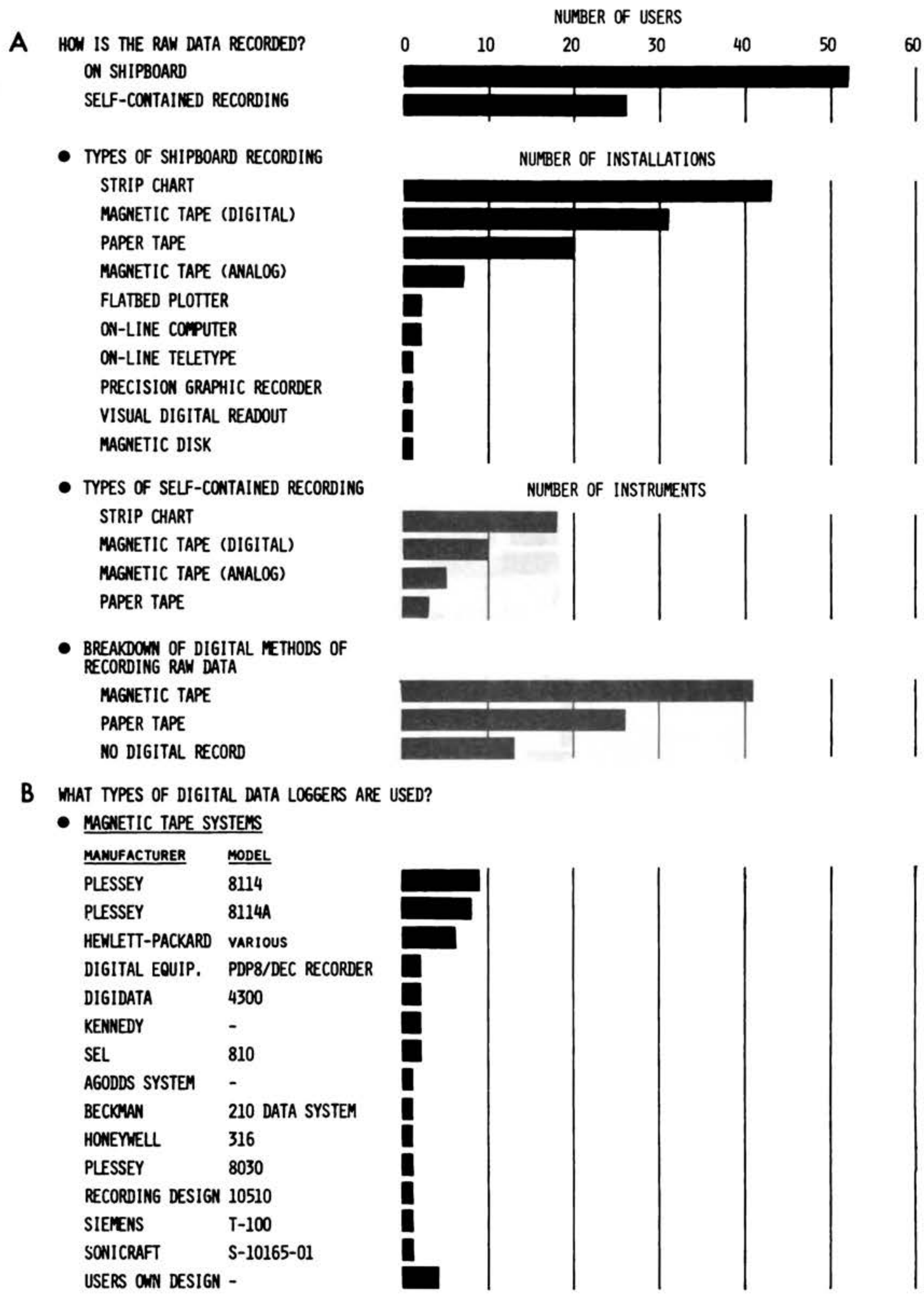


CHART 2 CONTINUED

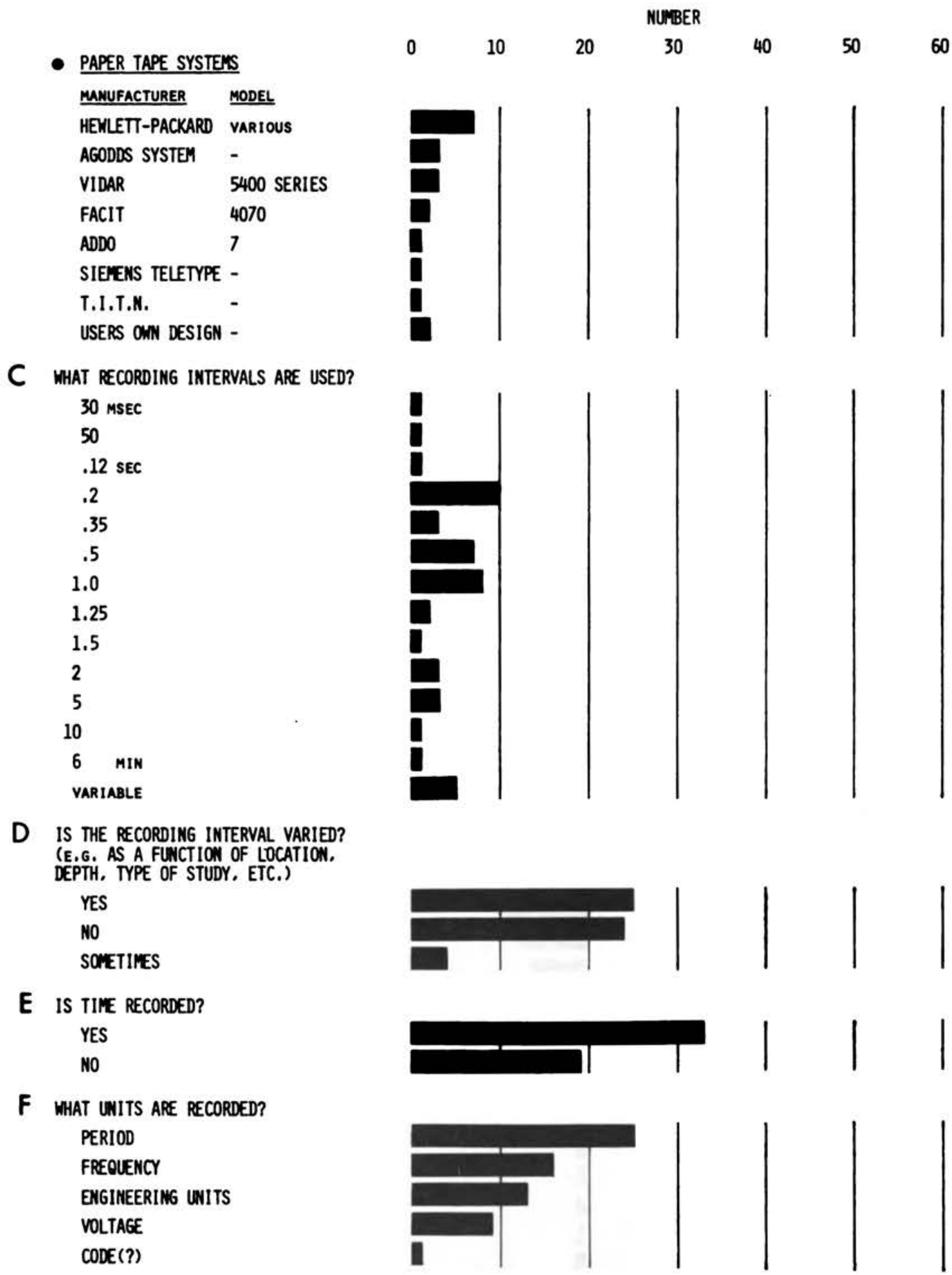


CHART 2 CONTINUED

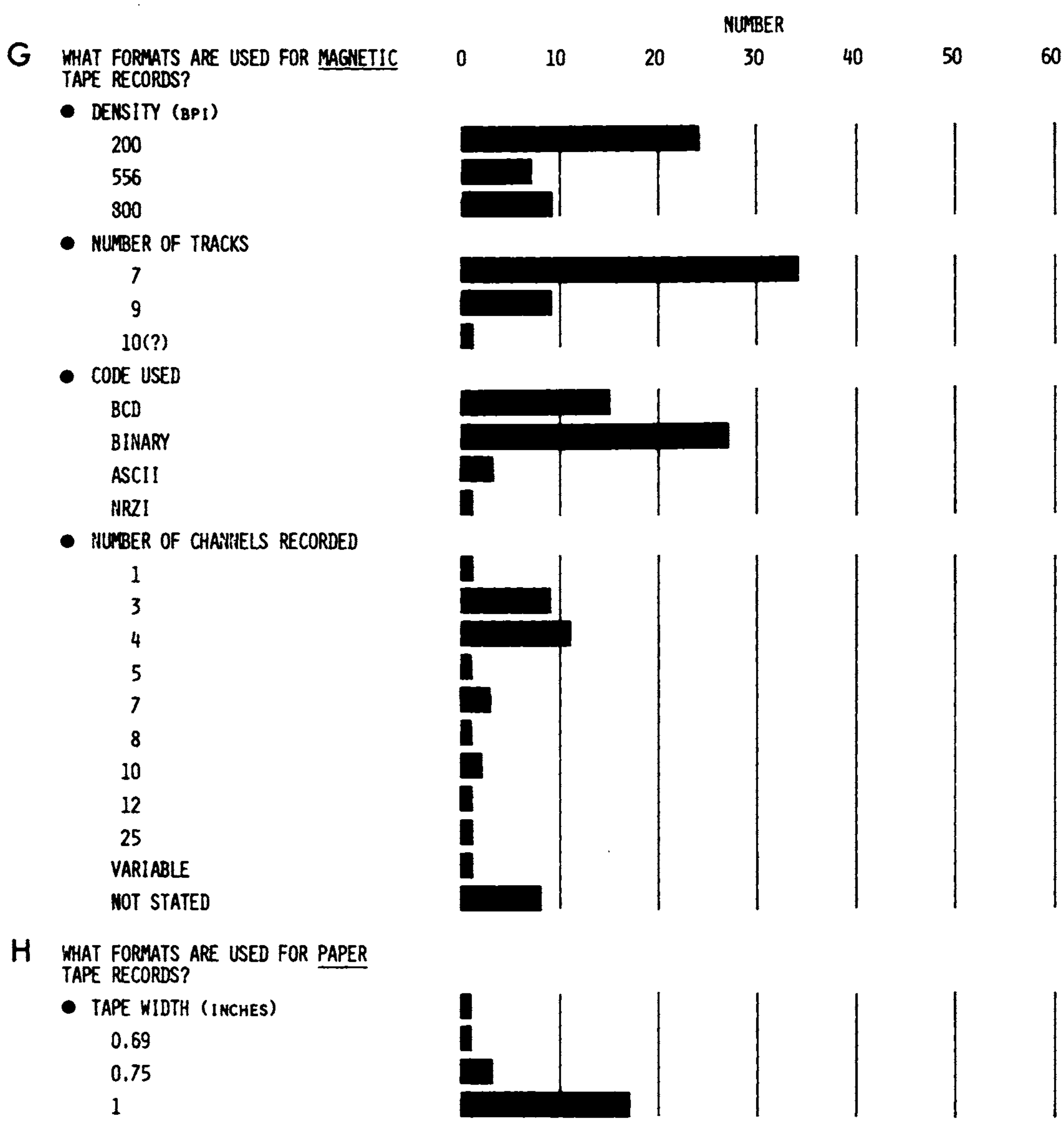
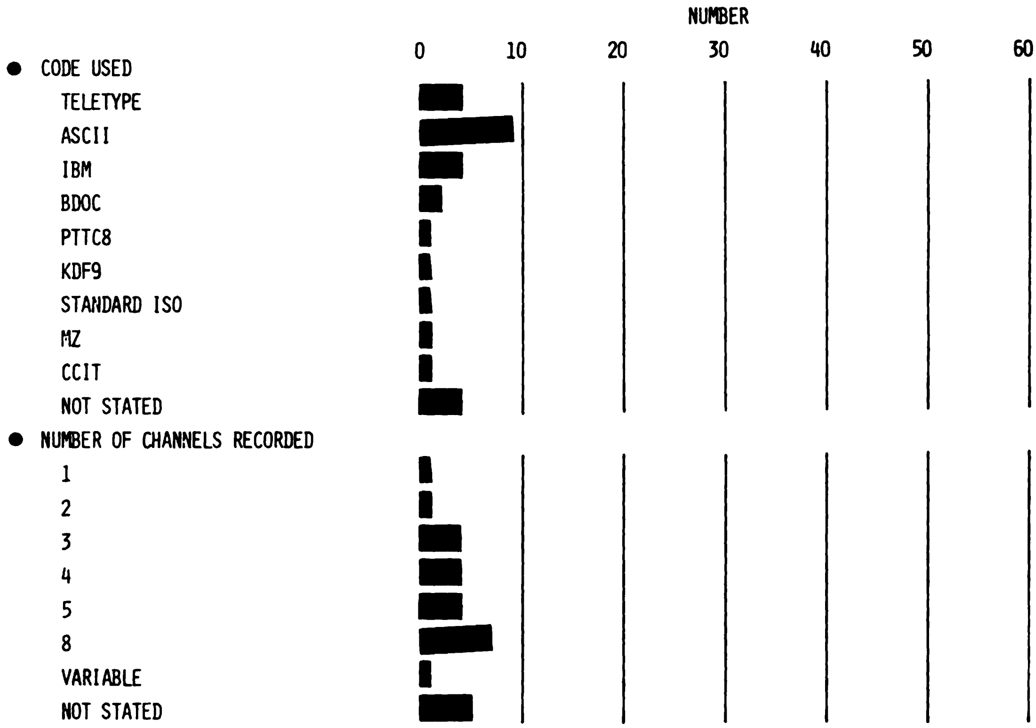


CHART 2 CONTINUED



NOTES TO CHART 2 - RAW DATA RECORDING

A. *Comparison between shipboard recording (data transmitted to shipboard, usually via a conducting seacable) and self-contained recording systems in the instrument housing. Most users record on shipboard using strip-chart recorders (usually an X-Y type recorder). Frequently, more than a single type of recording system is used with any one installation, i.e., strip-chart and magnetic tape.*

B. *Digital records are of greatest interest in dealing with the continuous nature of STD records. C through H examine the details of various systems and the methods of digital data logging. A great variety of magnetic and paper tape systems are in use, with Plessey being the most often used. Each system has its own recording format (record lengths, recording densities, etc.)*

C. *Generally, digital data is recorded sequentially (although all data channels may be scanned simultaneously) at a certain time interval per scan. Translated into data channels recorded per meter this represents a huge variation among users. Data channels recorded per meter is, of course, also dependent on the lowering rate of the instrument through the water column (CHART 4B, C) which varies tremendously.*

D. *Many users vary the recording interval depending on the type of study, depth, area of the study, etc., and most users vary the drop-rate of the STD as a function of the gradients encountered in the water column.*

E. *Shows how many users record time (from a digital clock) directly on their raw data tapes.*

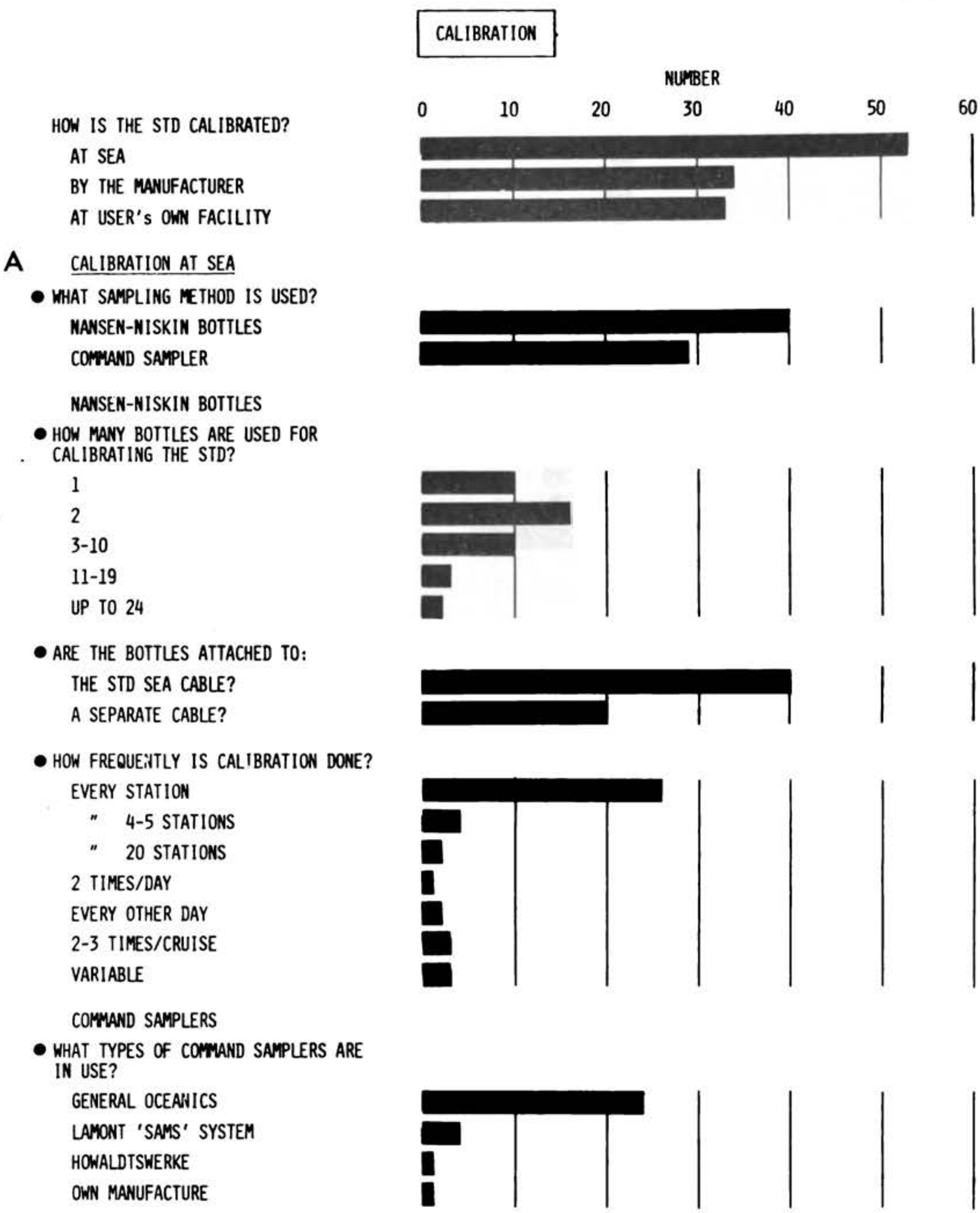
F. *Most systems that transmit data to recording equipment as an FM signal will record frequencies or periods on the raw data tapes which are converted to parameter ("engineering") units at a later stage in the data processing. Some digital data loggers convert to engineering units and record these directly.*

NOTES TO CHART 2 - RAW DATA RECORDING (continued)

G. *Formats used for magnetic tape records. Most users record their raw data on 7-track tape at 200 BPI in binary form with four data channels per scan. The tables show some considerable variation in recording formats, however.*

H. *Formats used for paper tape records. Most paper-tape users record their raw data on 1" tape in ASCII code, with 8 channels per scan.*

CHART 3


CALIBRATION
NUMBER
0
10
20
30
40
50
60
HOW IS THE STD CALIBRATED?
AT SEA
BY THE MANUFACTURER
AT USER's OWN FACILITY
A CALIBRATION AT SEA
● WHAT SAMPLING METHOD IS USED?
NANSEN-NISKIN BOTTLES
COMMAND SAMPLER
NANSEN-NISKIN BOTTLES
● HOW MANY BOTTLES ARE USED FOR CALIBRATING THE STD?
1
2
3-10
11-19
UP TO 24
● ARE THE BOTTLES ATTACHED TO:
THE STD SEA CABLE?
A SEPARATE CABLE?
● HOW FREQUENTLY IS CALIBRATION DONE?
EVERY STATION
" 4-5 STATIONS
" 20 STATIONS
2 TIMES/DAY
EVERY OTHER DAY
2-3 TIMES/CRUISE
VARIABLE
COMMAND SAMPLERS
● WHAT TYPES OF COMMAND SAMPLERS ARE IN USE?
GENERAL OCEANICS
LAMONT 'SAMS' SYSTEM
HOWALDTSWERKE
OWN MANUFACTURE

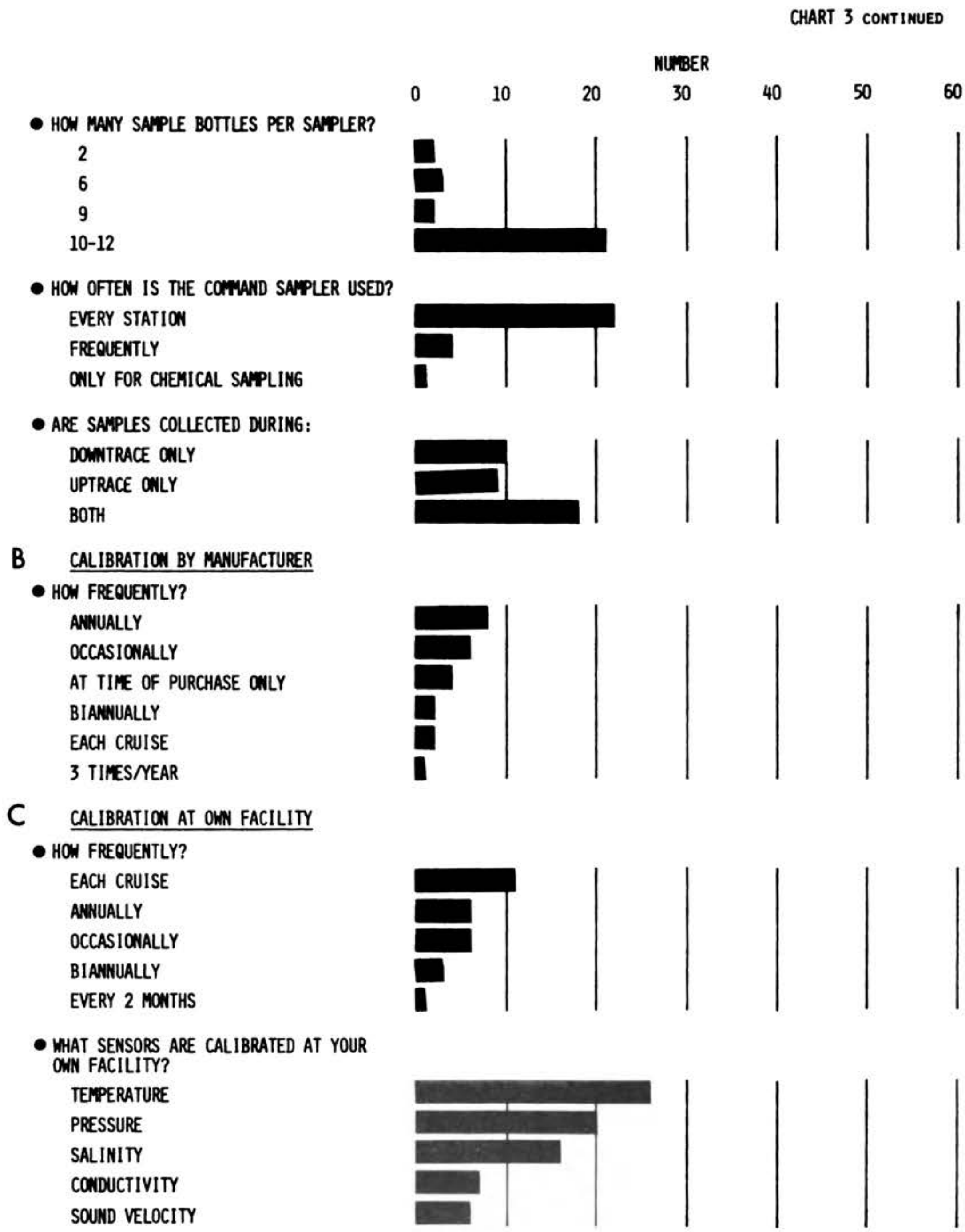


CHART 3 CONTINUED
NUMBER
0
10
20
30
40
50
60
● HOW MANY SAMPLE BOTTLES PER SAMPLER?
2
6
9
10-12
● HOW OFTEN IS THE COMMAND SAMPLER USED?
EVERY STATION
FREQUENTLY
ONLY FOR CHEMICAL SAMPLING
● ARE SAMPLES COLLECTED DURING:
DOWNTRACE ONLY
UPTRACE ONLY
BOTH
B CALIBRATION BY MANUFACTURER
● HOW FREQUENTLY?
ANNUALLY
OCCASIONALLY
AT TIME OF PURCHASE ONLY
BIANNUALLY
EACH CRUISE
3 TIMES/YEAR
C CALIBRATION AT OWN FACILITY
● HOW FREQUENTLY?
EACH CRUISE
ANNUALLY
OCCASIONALLY
BIANNUALLY
EVERY 2 MONTHS
● WHAT SENSORS ARE CALIBRATED AT YOUR OWN FACILITY?
TEMPERATURE
PRESSURE
SALINITY
CONDUCTIVITY
SOUND VELOCITY

NOTES TO CHART 3 - CALIBRATION

Three methods of calibrating the STD are in use: at-sea, at the user's own facility and by the manufacturer. Many users employ all three methods.

A. *<u>At-Sea Calibration</u>: "Calibration" is defined here to mean "standardization." By collecting samples of sea-water and using deep-sea reversing thermometers, comparisons can be made between the "classical" method of measuring the physical oceanographic parameters and the sensed output of the STD. The STD may then be corrected to the classical data obtained if statistically enough comparison points are made. This technique aids in keeping track of drifts and shifts in the instruments' output and most users agree that this is an essential part of STD data collection (CHART 6D). Two methods are used to perform the at-sea calibration: (1) Nansen or Niskin sampling bottles with reversing thermometers, attached to the STD sea cable or to a separate cable (e.g., doing a complete hydrographic station at the same location as the STD station); or (2) command samplers with the STD that collect a sample (after receiving command from the deck control equipment) in the same physical environment as the STD sensors while the STD output is being monitored for direct comparison.*

With Nansen or Niskin bottles, most users put 2 bottles (1 at the surface, 1 just above the STD) on the STD sea-cable at every station. With command samplers, 10-12 bottles are used on every station by the majority of users. Frequently additional sampling programs are carried out using command samplers to augment the STD program (CHART 4D). In using a command sampler there is a problem with whether to sample on the downtrace (shallow samples are taken down to deeper levels after collection and then brought back to the surface); uptrace (more comparable to standard hydrographic technique); or on both traces. Most people use samples collected during both descent and ascent of the STD. However, this raises the problems of which trace is used in the final data (CHART 5F) and whether uptrace (or downtrace) command samples are used to correct downtrace (or uptrace) STD data, etc.

NOTES TO CHART 3 - CALIBRATION (continued)

B. *Calibration by Manufacturer: Probably all STDs are calibrated in precise environmental tanks against known standards by the manufacturer before being sold. Most respondees to this question have their instruments recalibrated annually by the manufacturer.*

C. *Calibration at Own Facility: A large number of users have their STDs calibrated at frequent intervals at their own facilities.*

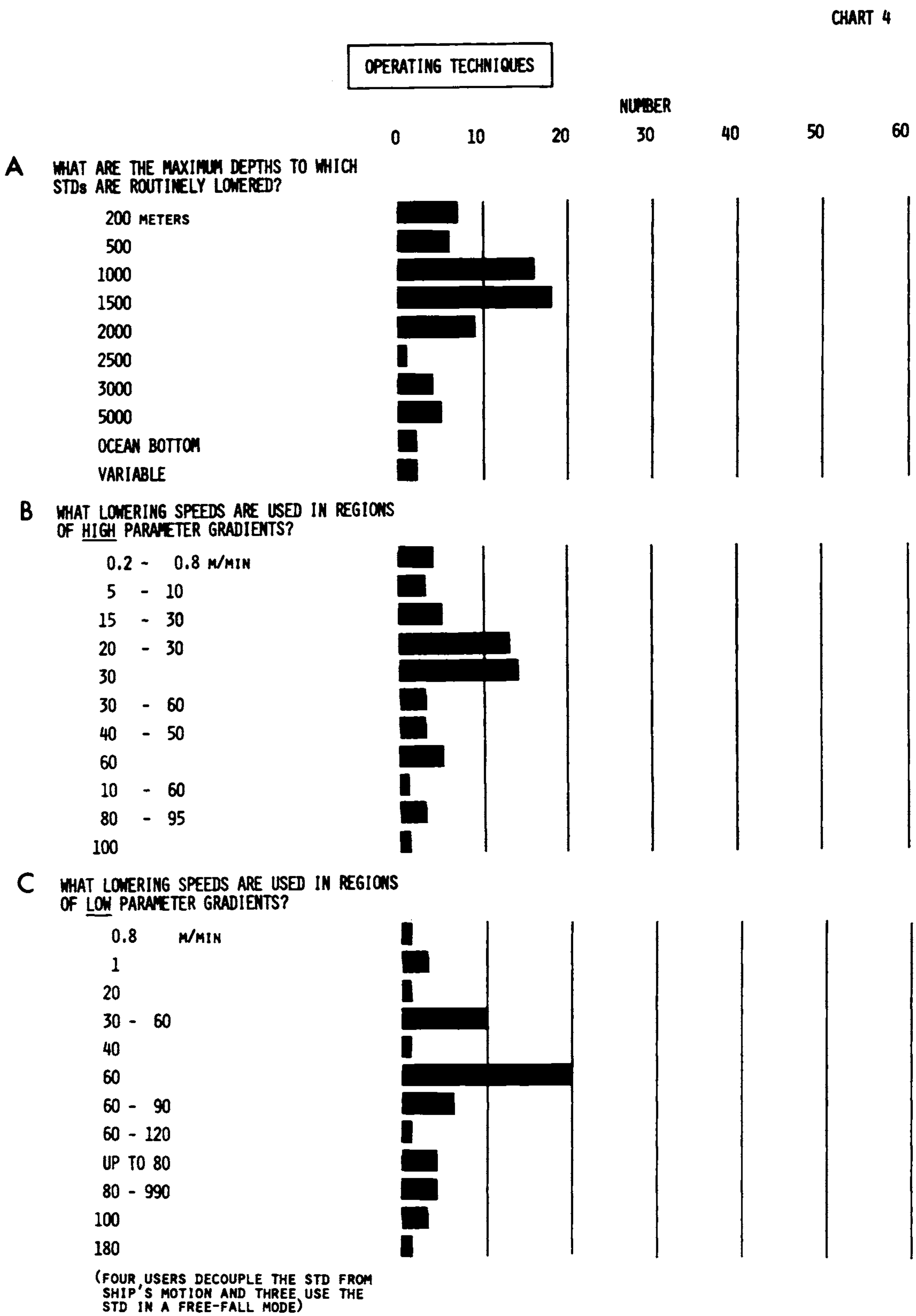
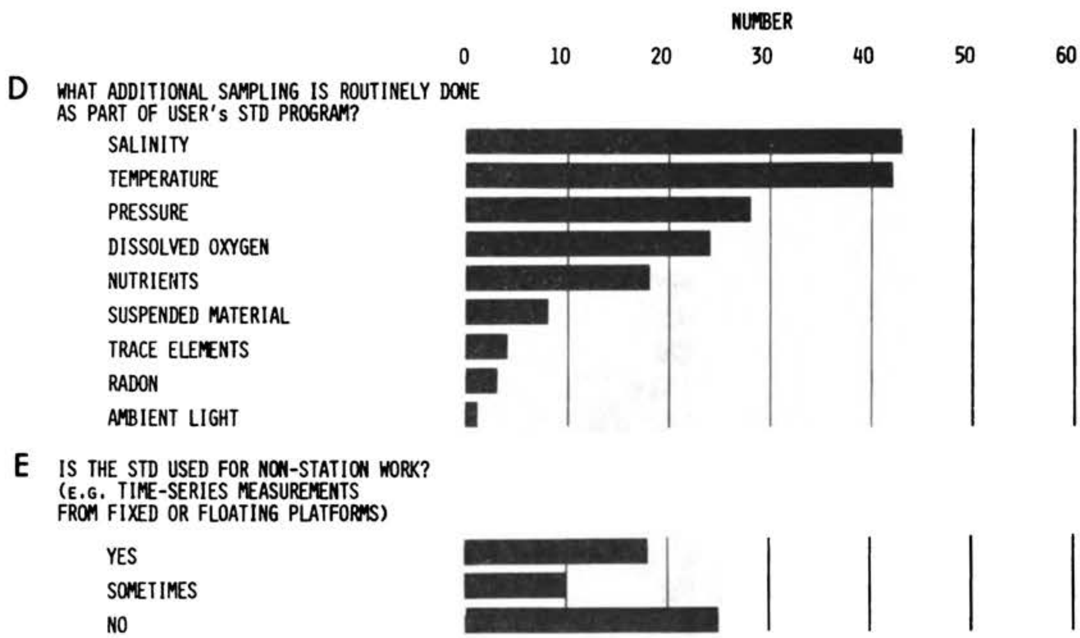


CHART 4
OPERATING TECHNIQUES
NUMBER
0
10
20
30
40
50
60
A WHAT ARE THE MAXIMUM DEPTHS TO WHICH STDs ARE ROUTINELY LOWERED?
200 METERS
500
1000
1500
2000
2500
3000
5000
OCEAN BOTTOM
VARIABLE
B WHAT LOWERING SPEEDS ARE USED IN REGIONS OF HIGH PARAMETER GRADIENTS?
0.2 - 0.8 M/MIN
5 - 10
15 - 30
20 - 30
30
30 - 60
40 - 50
60
10 - 60
80 - 95
100
C WHAT LOWERING SPEEDS ARE USED IN REGIONS OF LOW PARAMETER GRADIENTS?
0.8 M/MIN
1
20
30 - 60
40
60
60 - 90
60 - 120
UP TO 80
80 - 990
100
180
(FOUR USERS DECOUPLE THE STD FROM SHIP'S MOTION AND THREE USE THE STD IN A FREE-FALL MODE)

CHART 4 CONTINUED



NOTES TO CHART 4 - OPERATING TECHNIQUES

A. *Most of the STD users work only to intermediate depths ranging from 1000 to 1500 meters.*

B & C. *Because of the long time constants of most STD sensors, it has become practice to use a slower instrument lowering rate through regions of high parameter gradients (particularly temperature). An incredible range of drop-rates is used, varying from 0.2 to 100 m/min in regions of high parameter gradients (B), to 0.8 to 180 m/min in regions of low parameter gradients (C). The most frequent drop-rate employed is 30 m/min for (B) and 60 m/min for (C). To exemplify what the extremes mean in terms of station time, an 1800-m station would take 20 minutes (round-trip) at 180 m/min and 210 hours at 0.2 m/min.*

D. *Most people do some additional sampling as part of their STD program with salinity, temperature, pressure, dissolved oxygen and nutrients being the main parameters sampled.*

E. *Quite a few people use their STDs in other ways than vertical profiling from a ship on-station (e.g., time-series measurements, "yo-yoing" configurations), all of which present problems for archiving data collected in such ways.*

CHART 5

DATA REDUCTION, PROCESSING & STORAGE

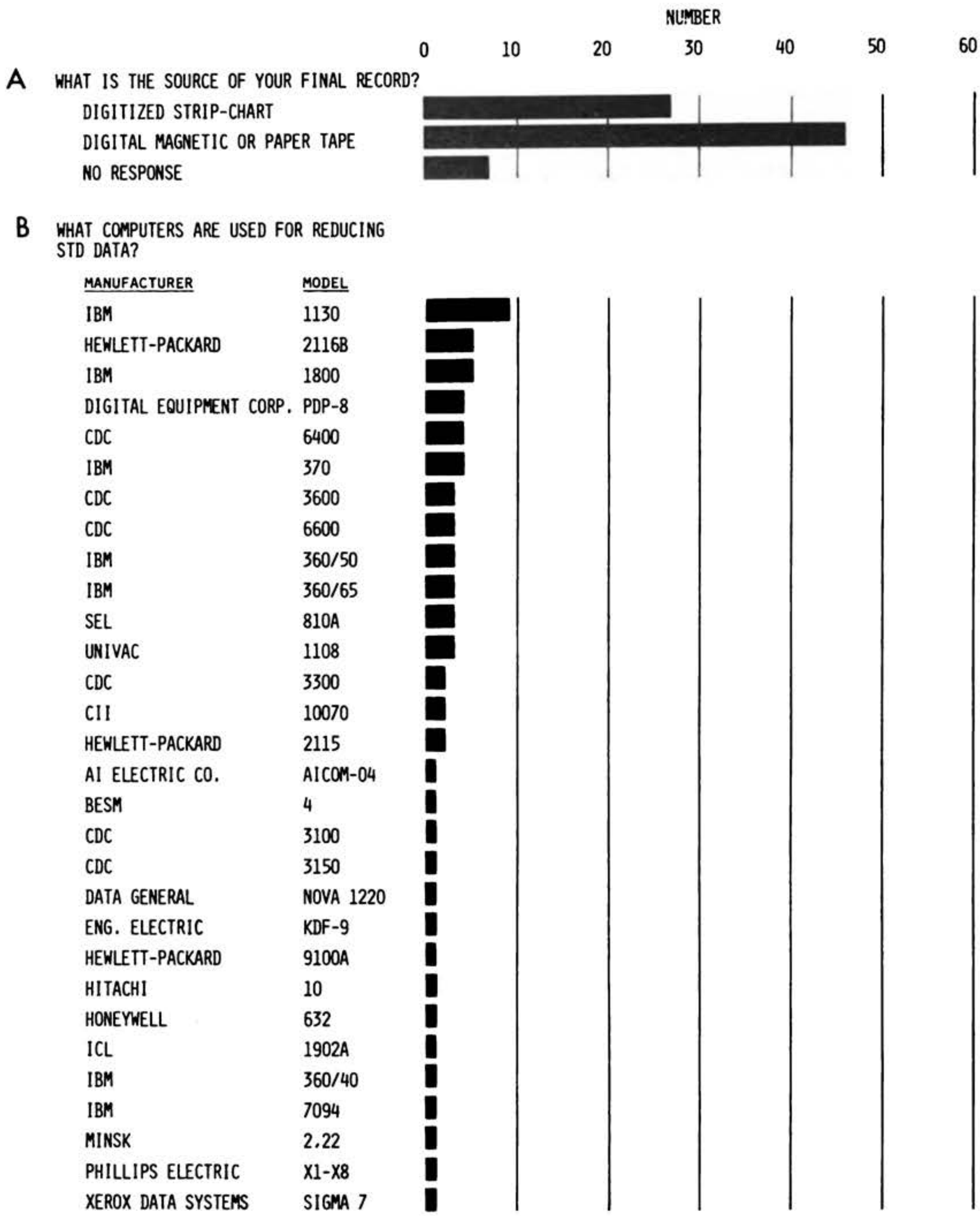


CHART 5 CONTINUED

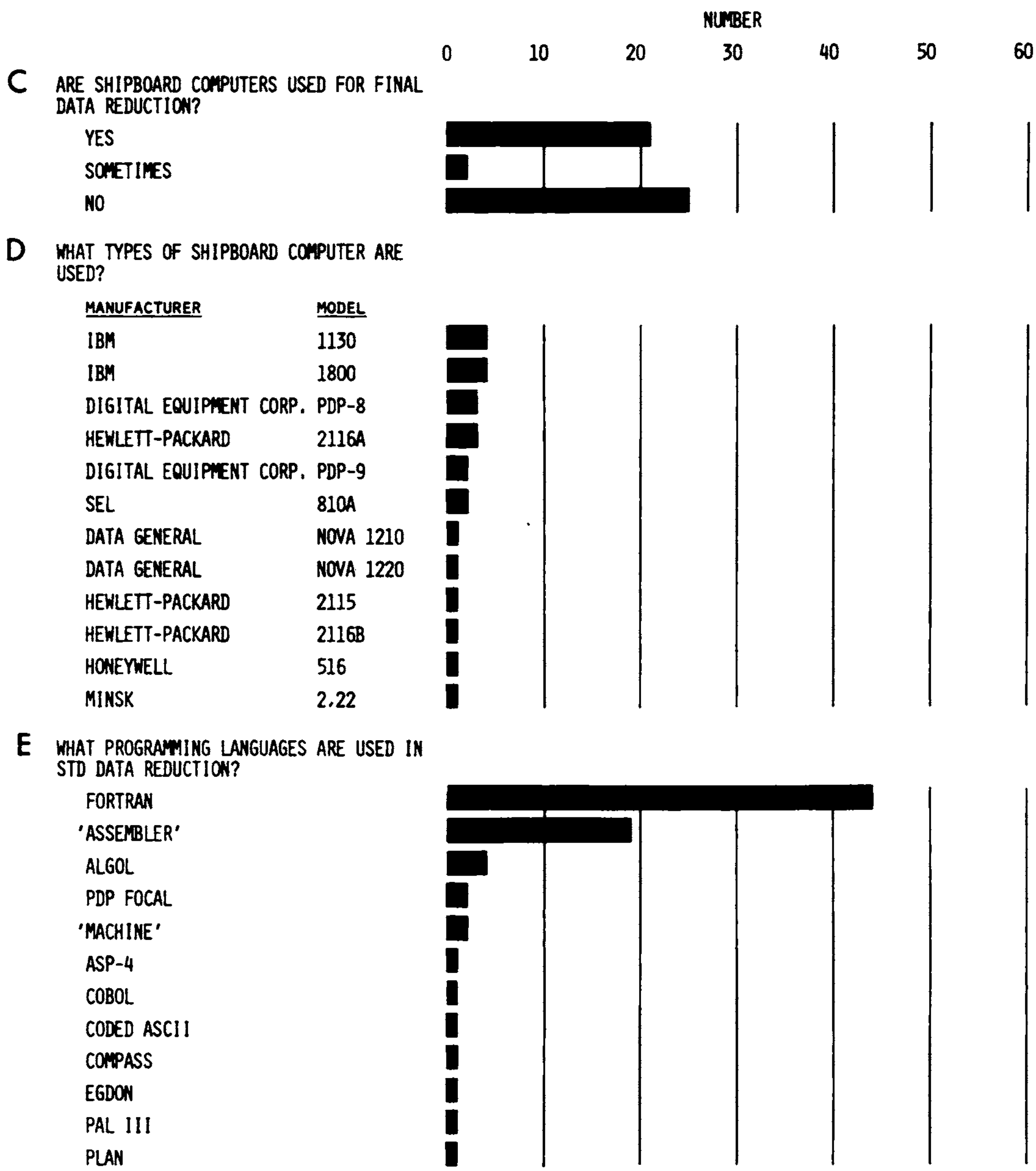


CHART 5 CONTINUED

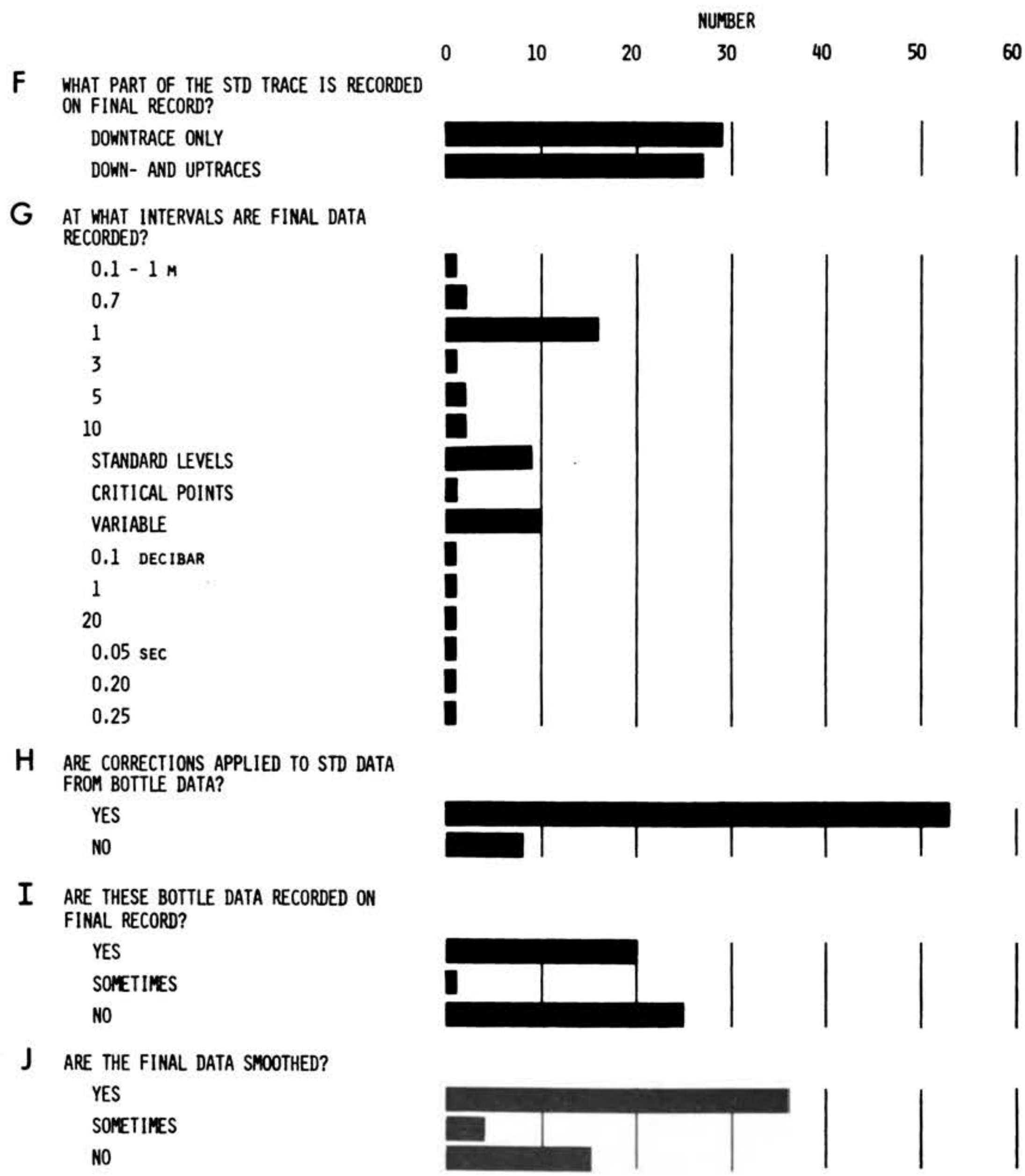
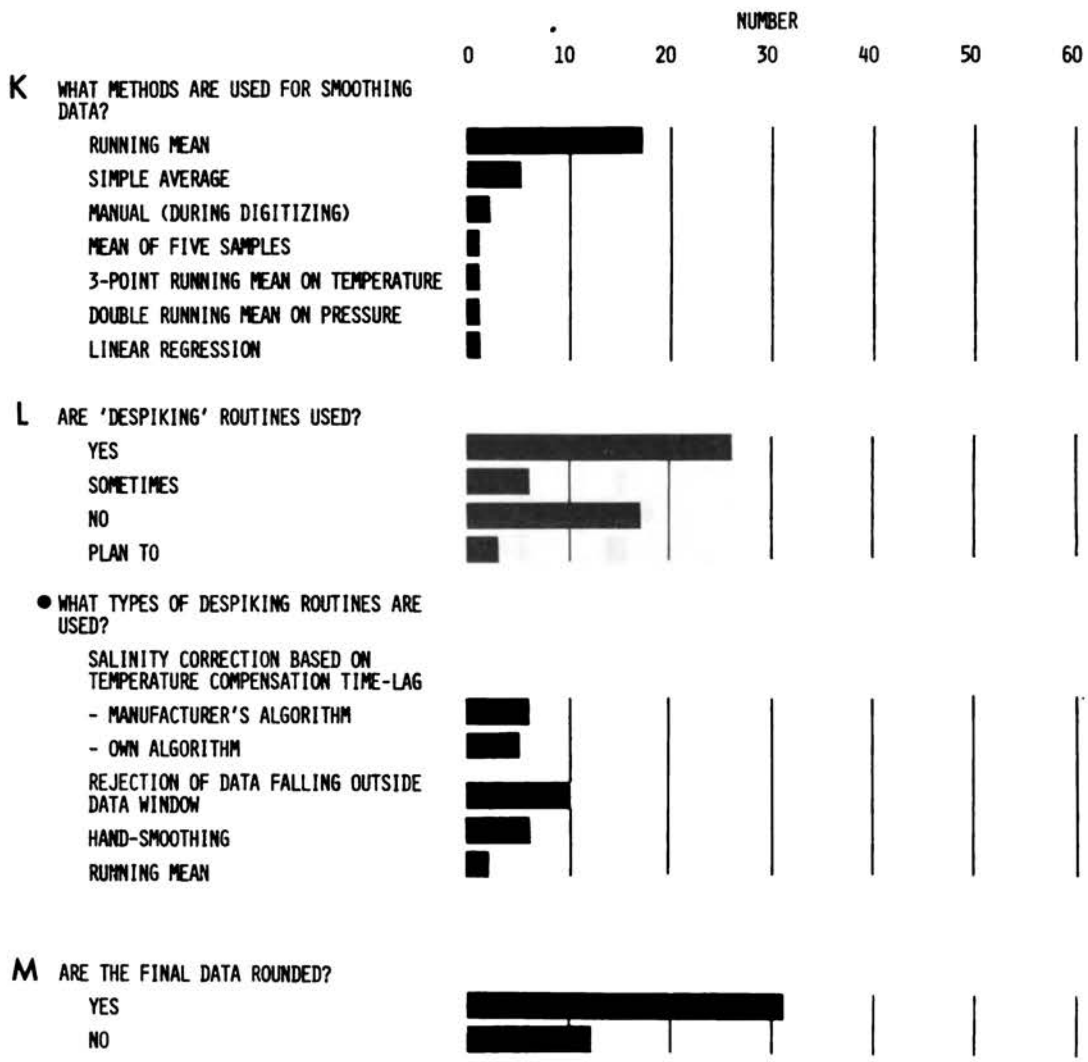


CHART 5 CONTINUED



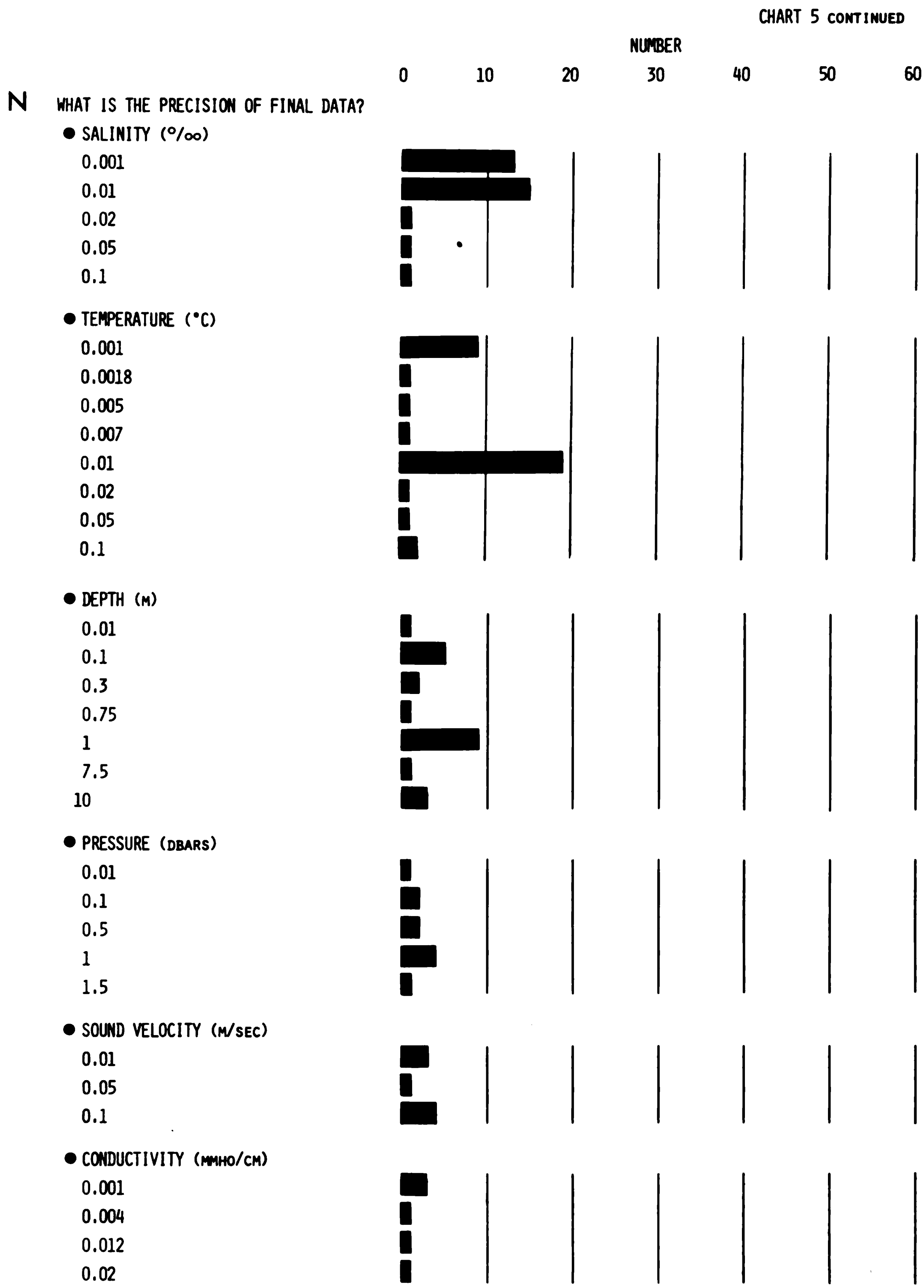
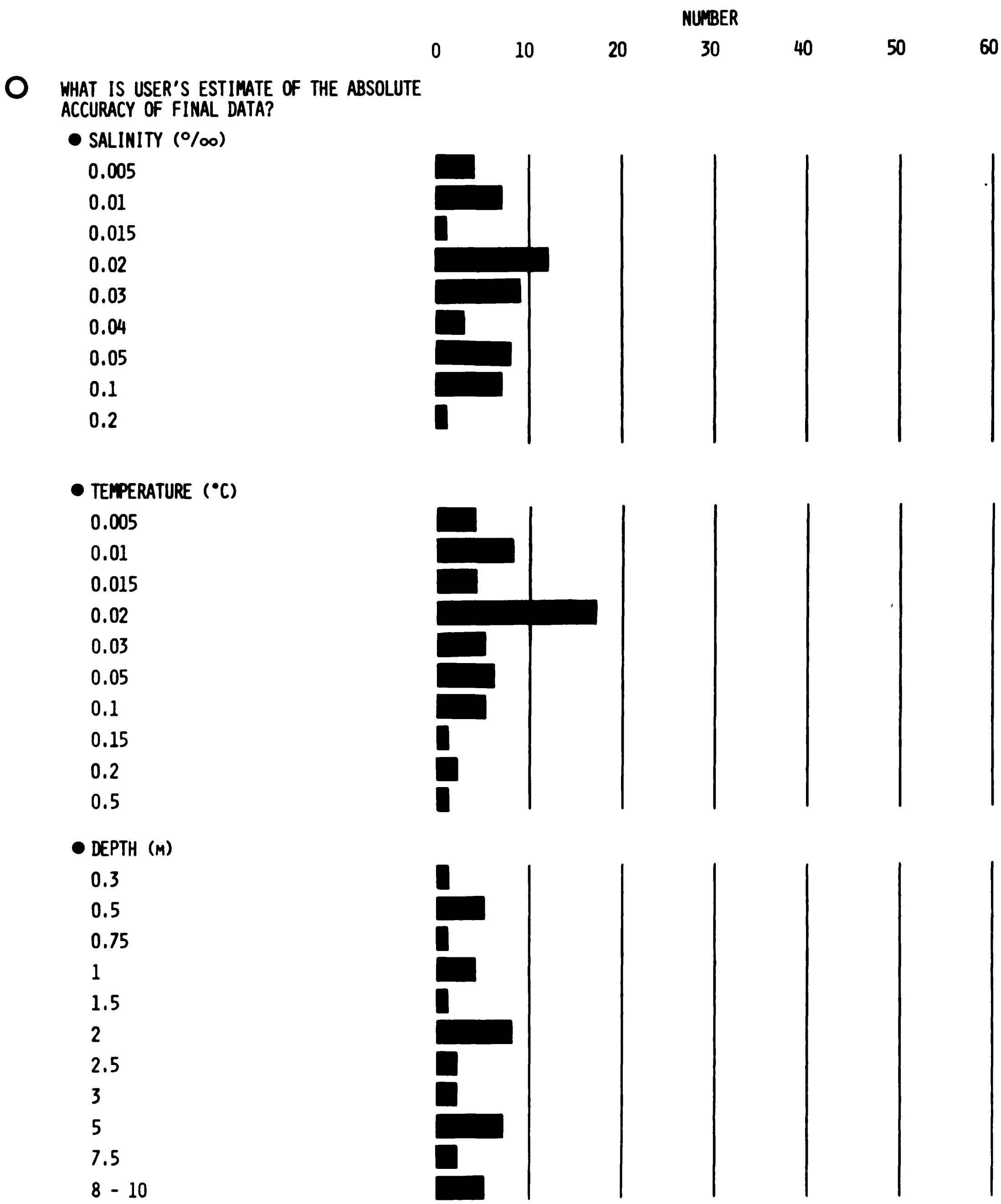


CHART 5 CONTINUED
NUMBER
0
10
20
30
40
50
60
N WHAT IS THE PRECISION OF FINAL DATA?
● SALINITY (°/oo)
0.001
0.01
0.02
0.05
0.1
● TEMPERATURE (°C)
0.001
0.0018
0.005
0.007
0.01
0.02
0.05
0.1
● DEPTH (M)
0.01
0.1
0.3
0.75
1
7.5
10
● PRESSURE (DBARS)
0.01
0.1
0.5
1
1.5
● SOUND VELOCITY (M/SEC)
0.01
0.05
0.1
● CONDUCTIVITY (MMHO/CM)
0.001
0.004
0.012
0.02

CHART 5 CONTINUED



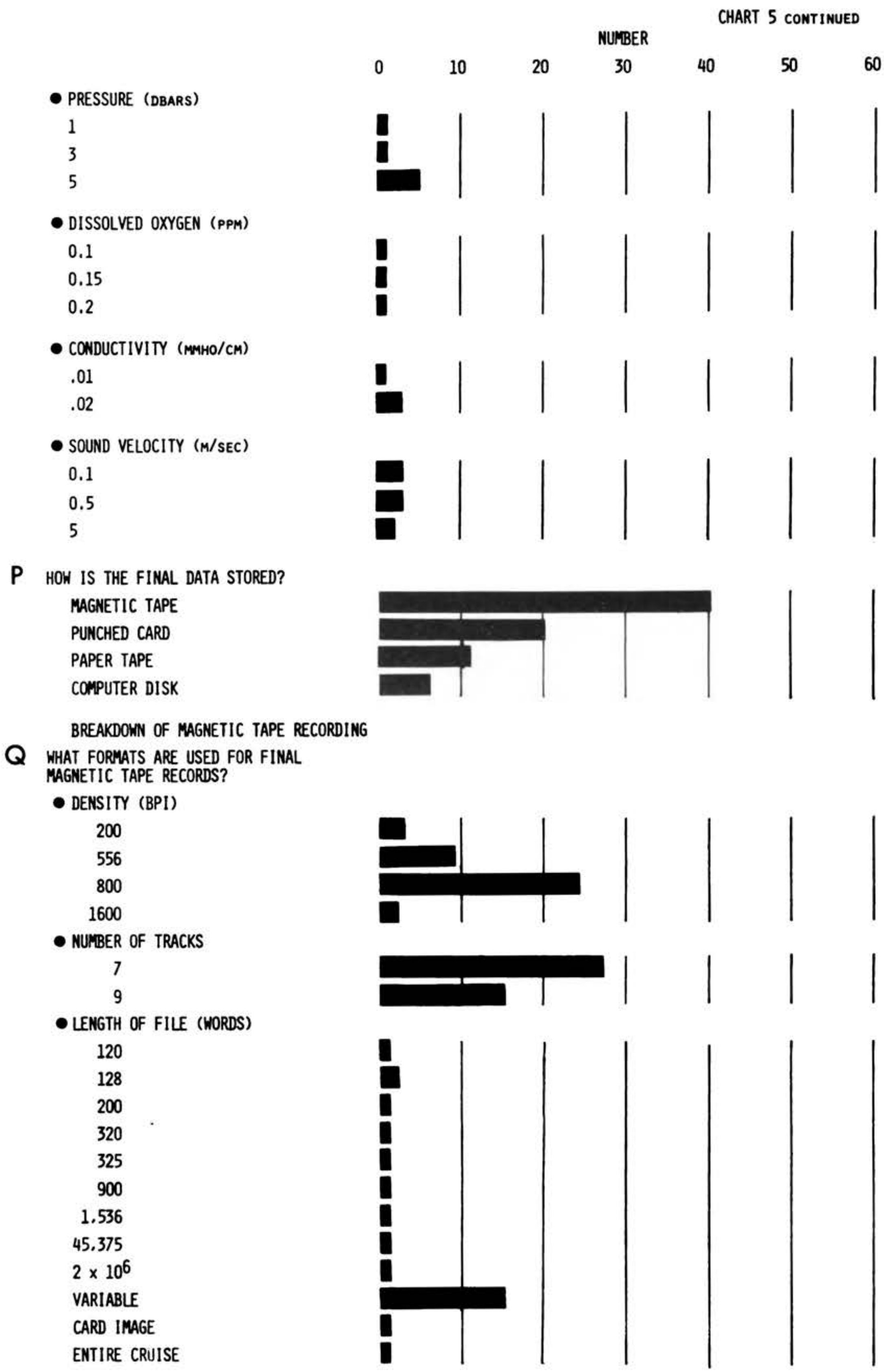


CHART 5 CONTINUED
NUMBER
0 10 20 30 40 50 60
● PRESSURE (DBARS)
1
3
5
● DISSOLVED OXYGEN (PPM)
0.1
0.15
0.2
● CONDUCTIVITY (MMHO/CM)
.01
.02
● SOUND VELOCITY (M/SEC)
0.1
0.5
5
P HOW IS THE FINAL DATA STORED?
MAGNETIC TAPE
PUNCHED CARD
PAPER TAPE
COMPUTER DISK
BREAKDOWN OF MAGNETIC TAPE RECORDING
Q WHAT FORMATS ARE USED FOR FINAL MAGNETIC TAPE RECORDS?
● DENSITY (BPI)
200
556
800
1600
● NUMBER OF TRACKS
7
9
● LENGTH OF FILE (WORDS)
120
128
200
320
325
900
1,536
45,375
2 x 10^6
VARIABLE
CARD IMAGE
ENTIRE CRUISE

CHART 5 CONTINUED

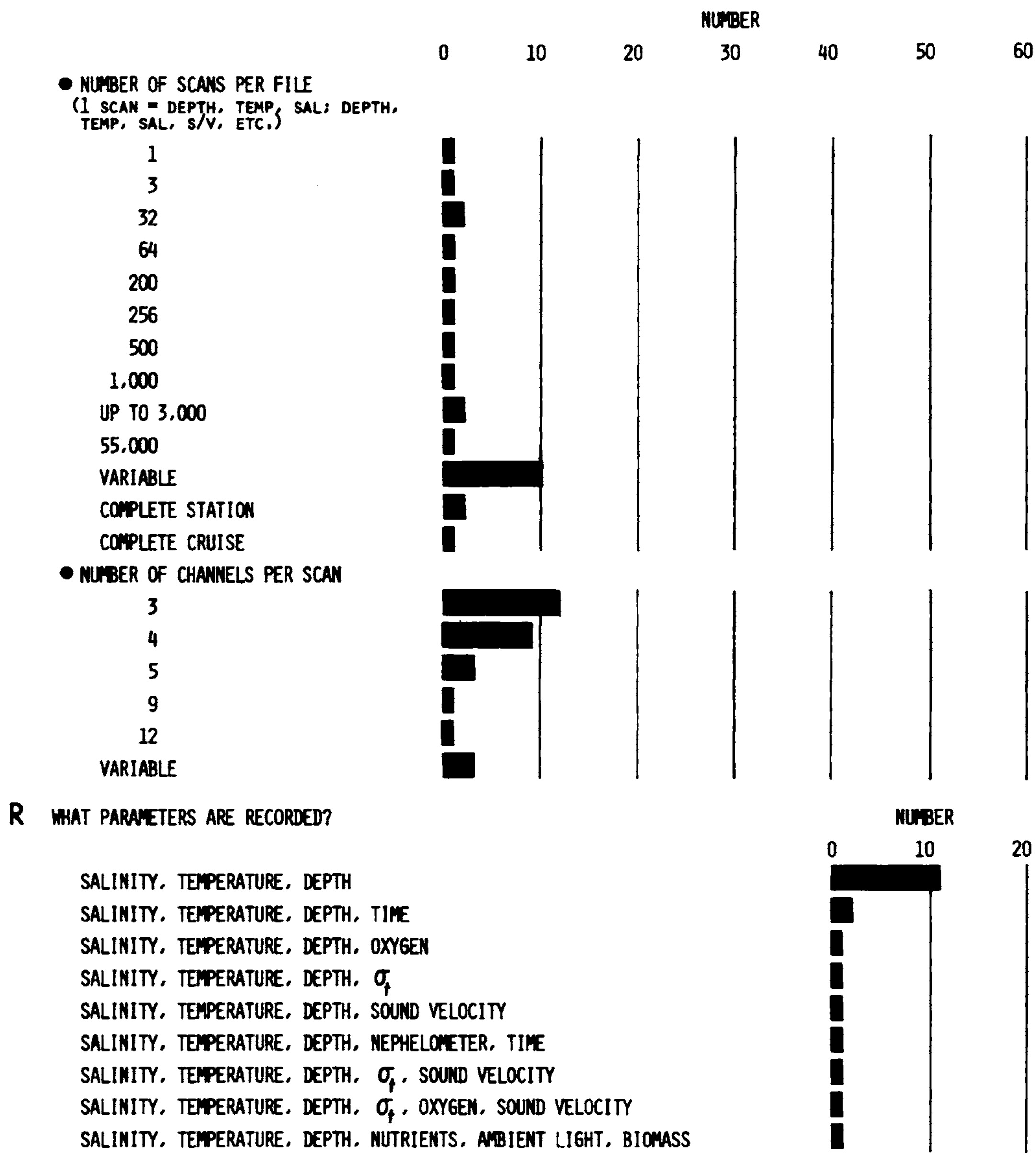
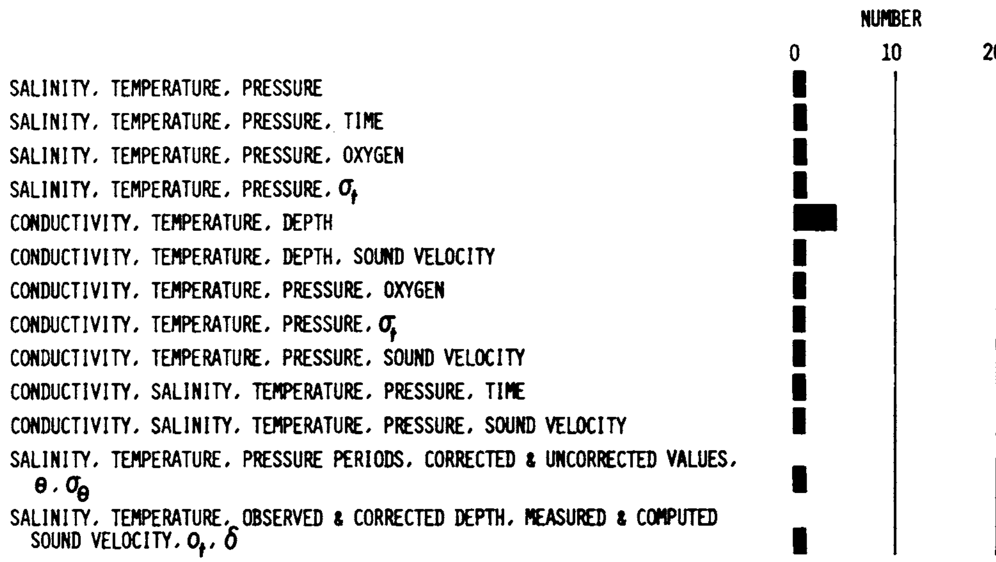


CHART 5 CONTINUED



NOTES TO CHART 5 - DATA REDUCTION, PROCESSING & STORAGE

A. *The final record of an STD station, suitable for submission to a data bank, will vary in quality depending on the source of the data. Users without shipboard or self-contained digital recording generally digitize their strip-chart analog records. Digital records can be operated on directly to perform the next stage of data reduction, normally a conversion to oceanographic units, followed by correction to at-sea or other calibration standards, smoothing and "despiking."*

B. *Not less than 30 different computers are used by STD people to reduce their data. The most frequently used computers are various models of IBM, CDC and Hewlett-Packard.*

C & D. *A number of users do their final data reduction aboard ship with "mini-computers" generally being the most popular for this job.*

E. *Fortran is the most frequently used programming language for STD data reduction.*

F. *Most users feel that the downtrace of an STD record is the most representative part of the record to use for their final data (quite frequently due to time considerations, the STD is turned off during the uptrace so that it can be raised at maximum speed). Many people use both downtraces and uptraces for their final record.*

G. *Final data is usually recorded in the form of temperature, salinity, etc., at certain depth increments, much like standard hydrographic-cast data. One level of data per meter is the most often used interval, but the range varies widely with some users reducing their data to standard levels and others to "observed" levels which will vary due to changes in droprate and recording rate.*

H. *Almost all users apply corrections to raw STD data from at-sea calibration data.*

NOTES TO CHART 5 - DATA REDUCTION, PROCESSING & STORAGE (continued)

I. *Less than half the responders record correction data on their final record - an important piece of information that is required to evaluate the veracity of the STD data (CHART 6D).*

J. *Smoothing techniques are frequently employed, a running mean being the most frequently used method. Various types of noise appear on STD records:*

(1) *transients (spikes) in the salinity trace caused by the long time constant of the compensating thermistor as the STD is lowered through a strong temperature gradient;*

(2) *various types of electronic noise caused by slip-rings, sea-cable, vibration, underwater connectors. These are occasional spikes that can occur anywhere in the record on any or all parameters, usually tending towards zero;*

(3) *noise introduced by inadequate digitizing techniques, e.g., counting too few periods of a sensor's output frequency. Many users either do not want all the data on their raw-data record in the final record, or want their final record in even increments of depth or pressure.*

K & L. *Despiking routines to eliminate noise from sources (1) and (2) are used by over half the responders. The routines used to-date for correcting for (1) are not accepted by all users and this problem constitutes one of the most serious types of error occurring in systems that compute salinity in situ (and to conductivity systems when salinity is to be computed for the final record). Several users employ a sliding data "window" which will correct for spikes of type (2), but this too is difficult to apply as genuine gradients in the ocean may well be as great as the "gradients" of the spikes. Hand-smoothing is often employed, particularly in digitizing analog strip-chart records.*

NOTES TO CHART 5 - DATA REDUCTION, PROCESSING & STORAGE (continued)

M. *Most users round-off their data to a significant decimal digit for presentation of their final record.*

N. *The precision of the final data varies considerably among users - somewhat reflecting the precision of the particular instrument used to generate the data, the number of significant digits to which the data is rounded, and the limitations of digitization techniques (e.g., a 0-6000 m depth sensor digitized to 1 part in 10^4 will have a precision of 0.6 m no matter how finely the data is sampled).*

O. *The users estimate of the absolute accuracy of his final data is of great interest to NODC to evaluate and code archived data. The shape of the histograms show somewhat less than Gaussian distribution. For salinity, the very wide "skirts" of the histogram curve reflect the general lack of confidence in STD salinity data. The slightly better shape of the temperature histogram shows a distinct peak at $\pm 0.02^{o}C$, while the complete lack of shape to the depth histogram is caused by the large variation in ranges covered by different users' pressure sensors. Of course, all these curves are affected by some STD systems being, by design, less accurate than others.*

P. *The method of storing final data by the user is of prime interest to NODC in converting STD data to a standard format for archiving. Most people use digital magnetic tape as the physical mode of storage. But, approximately one-half again use punched card data which form represents about the easiest way of transferring data from user to NODC, in a format sense, although many hundreds of punched cards may be necessary for each STD station. NODC prefers magnetic tape to punch cards. Most of the questions here were directed to magnetic-tape users although quite a few people use paper tape.*

NOTES TO CHART 5 - DATA REDUCTION, PROCESSING & STORAGE (continued)

Q. *Magnetic-tape formats: most data is stored on 7-track tape at 800 BPI, but an incredible variety of methods is employed for arranging files, channels of data, stations, and cruises within the data tape. The length of file which (unless the question has been misinterpreted) would create some difficulties for NODC if it exceeds 35k bytes in length. Within each file of data there may be several scans (depth intervals) of STD information, i.e., a user may record 32 levels of depth, temperature, salinity, etc., within one file. The answers varied from 1 to 55,000, to one station to one cruise per file. Again, the majority of people said the number of scans per file was variable. Within each scan there will be several channels (e.g., depth, temperature, salinity, would be 3 channels). This varied from 3-12 channels, with 3 and 4 being most often used.*

R. *A great variety of parameters is recorded by different users on their final record, with salinity, temperature, and depth being by far the most frequently recorded.*

CHART 6

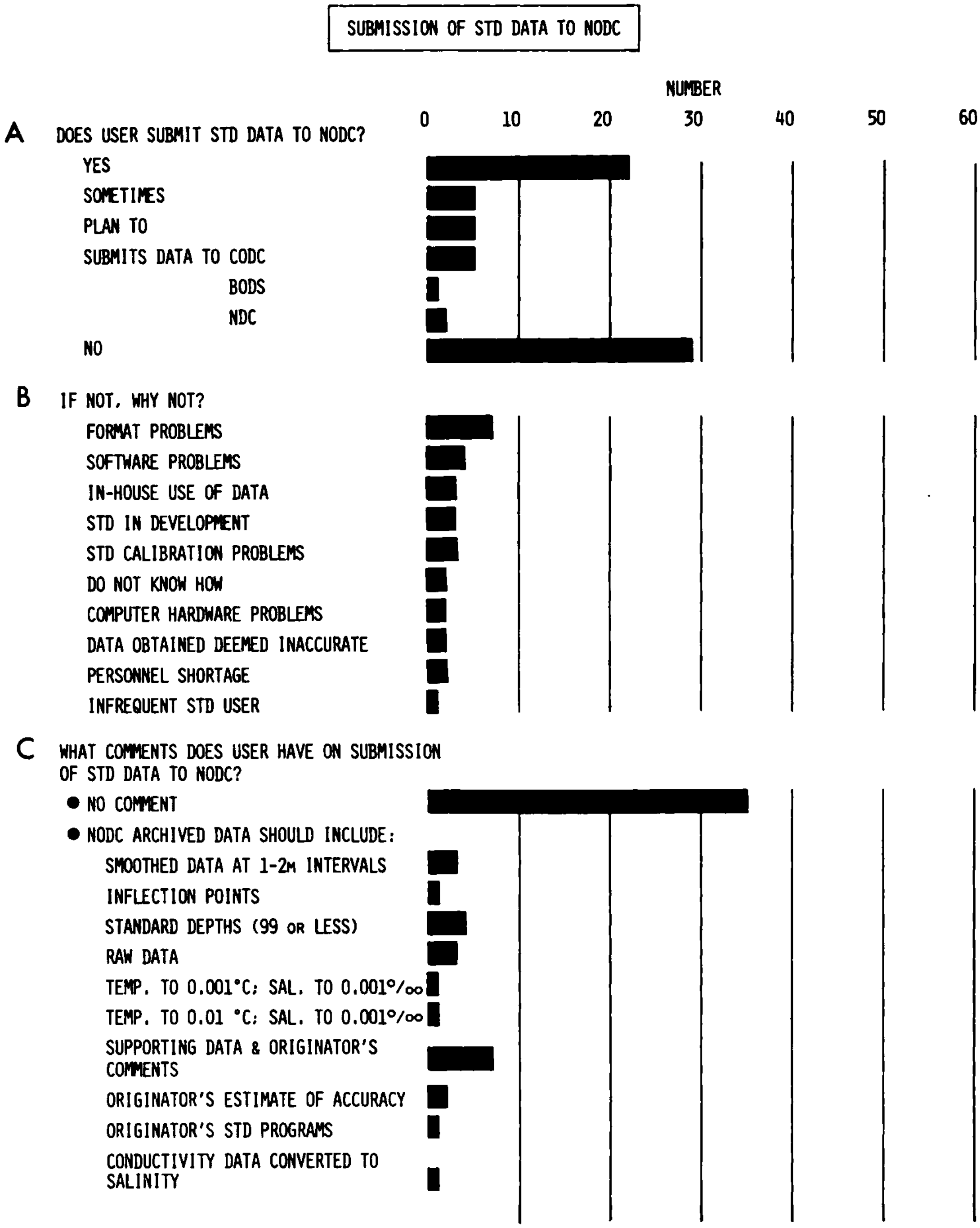
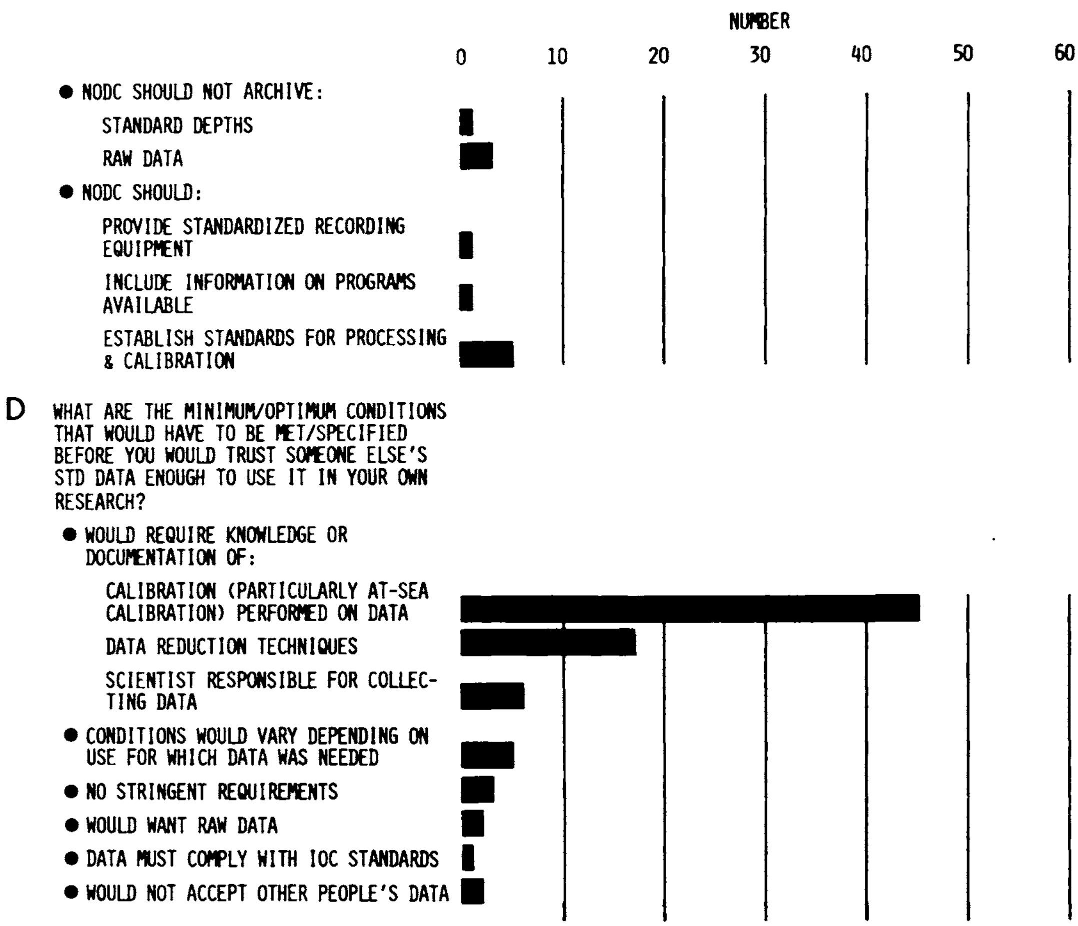

SUBMISSION OF STD DATA TO NODC
NUMBER
0
10
20
30
40
50
60
A DOES USER SUBMIT STD DATA TO NODC?
YES
SOMETIMES
PLAN TO
SUBMITS DATA TO CODC
BODS
NDC
NO
B IF NOT, WHY NOT?
FORMAT PROBLEMS
SOFTWARE PROBLEMS
IN-HOUSE USE OF DATA
STD IN DEVELOPMENT
STD CALIBRATION PROBLEMS
DO NOT KNOW HOW
COMPUTER HARDWARE PROBLEMS
DATA OBTAINED DEEMED INACCURATE
PERSONNEL SHORTAGE
INFREQUENT STD USER
C WHAT COMMENTS DOES USER HAVE ON SUBMISSION OF STD DATA TO NODC?
● NO COMMENT
● NODC ARCHIVED DATA SHOULD INCLUDE:
SMOOTHED DATA AT 1-2M INTERVALS
INFLECTION POINTS
STANDARD DEPTHS (99 OR LESS)
RAW DATA
TEMP. TO 0.001°C; SAL. TO 0.001°/oo
TEMP. TO 0.01 °C; SAL. TO 0.001°/oo
SUPPORTING DATA & ORIGINATOR'S COMMENTS
ORIGINATOR'S ESTIMATE OF ACCURACY
ORIGINATOR'S STD PROGRAMS
CONDUCTIVITY DATA CONVERTED TO SALINITY

CHART 6 CONTINUED



NOTES TO CHART 6 - SUBMISSION OF STD DATA TO NODC

A. *Presently, the majority of STD users do not submit STD data to NODC. Foreign users submit data to their own data centers (CODC = Canadian Oceanographic Data Centre, BODS = British Oceanographic Data Service, NDC = DOD (National Data Centre, Hamburg).*

B. *Users which did not submit their data to NODC, explained that it was mostly due to computer software problems, lack of knowledge, funds, or capable personnel to put the data into some form suitable for submission.*

C. *When asked for general comments on submission of STD data to NODC, most users declined to comment. Of those that commented, there seems to be a diversity of opinion as to how and what data should be stored at NODC and what NODC should provide to aid the STD user in submitting his data.*

D. *The question, "What are the minimum/optimum conditions that would have to be met/specified before you would trust someone else's STD data enough to use it in your own research?" produced the overwhelming response that information on what calibration, particularly at-sea calibration, was performed on the data would be necessary before other people's data would be used. This was followed by inquiries about what data reduction techniques were used and which scientists were responsible for collecting the data. Two responders said they would not accept other people's STD data at all.*

CHAPTER 3
SURVEY OF STD SECONDARY USERS

A number of potential secondary users of STD were asked about their requirements for STD data. Nearly all reported that the quality of the STD data strongly depends on the collector.

> **Credibility may be characterized by the reputation of the experimentalist in the scientific community, by the reasonableness of the experimental set-up, by the reputation of the primary sensors used and by the reasonableness of the preprocessing techniques applied to the acquired data. Documentation concerning each of these four points must accompany the data.*

Documentation is always necessary in any data collection program for proper evaluation of data, but it is especially true when collection and processing methods vary such as with STD data.

Many of the secondary users want an NODC or collector's evaluation of the data.

> **I might suggest that along with the digital file containing the station information you include a quality index: on a scale of 0 to 5 for instance. Thus, the researcher would have some idea of how well the STD matched or was "massaged" to coincide with independent observations at various depth intervals (i.e., a Niskin sampler), just at the surface or not at all.*

---

* Individual replies to the questionnaire

However, the basic information to permit data evaluation should be supplied to the secondary user.

> **It is that user's responsibility to apply further filtering with regard to time and space (structure) intervals consistent with his own analysis. The use of the classic depth spacing of 10, 20, 30, 50, 75, 100, etc. meters as standards seems to be irrelevant to STD observations.*

The ability of the STD to define features of little vertical extent in the water column which may be significant processes on a wide range of scales requires special attention by NODC. Also, NODC must permit much flexibility in the formats used in supplying STD data.

> **Because it is sometimes hard to decide which maxima and minima are major (the decision will vary from problem to problem), we should be able to call out data digitized at quite close intervals - down to a meter, I suppose, in the good records. But printing these out and looking at the columns is expensive and time-consuming. This is where a xerox of the individual records would be useful. This may be more expensive than tape but may serve far better for searching.*

Some general remarks that were made regarding the needs of the secondary user follow.

Many of the secondary users of STD data require equivalent series-cast data, as may be needed in the case of preparation of standard oceanographic atlases or for large-scale water-mass distribution studies. The standard-level listing would be satisfactory. Some users, however, may use the STD's special capability by requiring data points at maxima or minima in the temperature and salinity distribution within the water column, by noting the depth of abrupt gradient changes, or noting the thickness of various layers. In this case, the secondary user requires the standard-level data points to be expanded to include not only the standard-level data but also extrema, inflexion, and gradient change data points. Such an expanded

---

* Individual replies to the questionnaire

standard-level data listing will meet the requirements of the great majority of secondary users.

Many of the primary collectors of STD data have as their research objective the study of the fine structures and mixing processes rather than larger-scale thermohaline structure. The data are carefully collected to meet the requirements of the collectors. The secondary user interested in such studies usually contacts the collector directly to learn about the details of the data collection and processing. The accumulation of high quality STD data which could be used for fine structure work suggests that NODC should archive and supply the secondary user with closely spaced data points for such data. When such data is used by a secondary user, heavy documentation with regard to methods of collection and processing are necessary. The number of secondary users requiring closely spaced STD data may at first be limited but as STD operations become more routine, and the records become more noise-free, the use of closely spaced data-point listings will increase.

The secondary user requires data in a number of formats. Magnetic tape or cards with a listing is generally required. In the case of closely spaced data points, an analog trace would be desirable. A cathode-ray tube display capability would be useful.

Chapter 4
RECOMMENDATIONS

This chapter outlines specific recommendations to NODC with regards to: techniques which yield results suitable for secondary users; interfiling of data collected using different techniques and recording systems; filtering and/or compression methods to be applied after receipt of STD data; and, information needed by secondary users of STD data regarding data collection methods.

RECOMMENDED TECHNIQUES WHICH YIELD RESULTS SUITABLE FOR SECONDARY USERS

At-Sea Standardization - For at-sea standardization, it is recommended that:

1. Sufficient classical data points be collected at sea with all STD stations to achieve Nansen-bottle accuracy at all points. This should require at least 6 (but preferably 10) classical data points throughout the water column for a deep (surface to 4000 m) STD station; and 2 or 3 data points for a shallow (less than 1500 m) STD station.

2. Classical data points be taken in close proximity to the STD sensor package so that the effects of temperature and salinity gradients are slight when compared to the accuracy of the classical data points. This is best achieved in homogeneous zones, where the ship's motion will not cause large variations in STD output when transmitted to the sensor package during collection of bottle and thermometer data.

3. Classical data points be obtained at a sufficient number of depth levels to define correction curves which are often non-linear (Figure 1).

4. Each STD station, if possible, be standardized to correct for temporal drifts of sensor output.

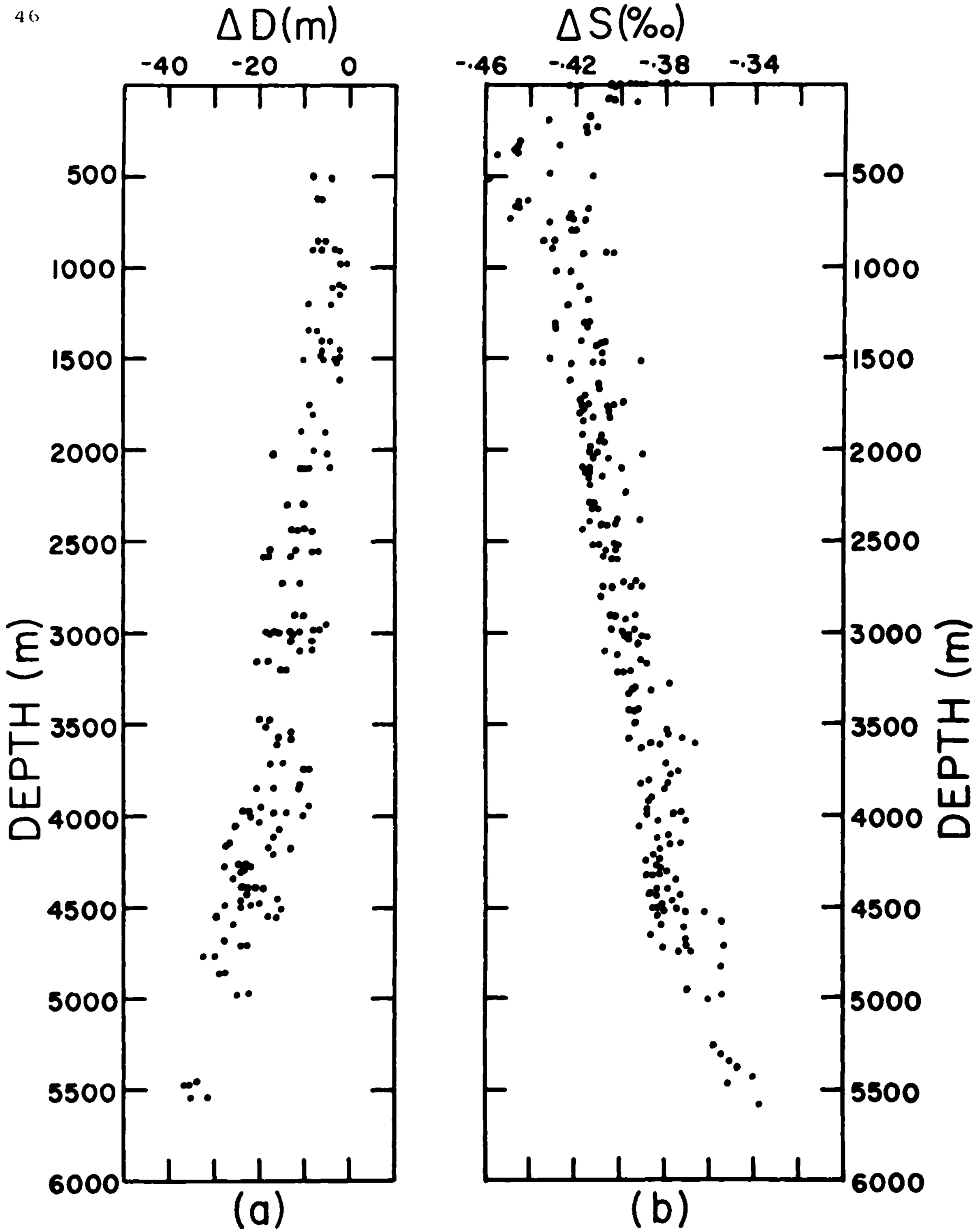


Figure 1. STD calibration data from 40 stations taken on R/V ATLANTIS II Cruise 68, 1972. (a) Difference between thermometric depth from SAMS thermometers and STD depth from manufacturer's linear depth/frequency relationship (ΔD) plotted as a function of thermometric depth. (b) Difference between SAMS salinities (determined with lab salinometer) and STD salinity from manufacturer's linear salinity/frequency relationship (ΔS), plotted as a function of STD depth.

5. Classical data points used in standardization be obtained during the profile which is being standardized, e.g., data points collected during the descent of the sensor package through the water column (descent profile) be used to correct the descent profile.

Data Collection Techniques - For data collection techniques, it is recommended that:

1. The lowering rate of the STD sensor package be slow enough to enable definition of structures in the order of meters. This is dependent on the time constants of the sensors, the motion of the ship, and the thermohaline gradients. However, the lowering rate must not be so slow as to degrade the resolution by poor flushing of the conductivity sensor.

2. As a general rule, a rate of 20 to 30 m/min be used for high gradients (as thermocline zones), 40 to 60 m/min for moderate gradients, and 60 to 80 m/min for low gradients (the desirable lowering rate will depend somewhat on the type of sensor used).

INTERFILING OF DATA COLLECTED USING DIFFERENT TECHNIQUES AND RECORDING SYSTEMS

STD Data Files at NODC - It is recommended that a file of raw data from the originators' digital data acquisition systems not be maintained at NODC, but that files of STD data that fall into the following categories be archived at NODC:

1. Ultra-High-Resolution Data File. This is a specialized category which contains data collected by microstructure probes that record on digital magnetic tape with data intervals less than 1 meter. Since the usefullness of such a file is not clear and the potential size of storage facilites is so great, it is recommended that only a listing of the availability of such a file, by the originator, be maintained at NODC.

2. High-Resolution Data File. Data falling in this category is obtained from STD systems that record eith on-deck or internally on digital magnetic- or paper-tape (preferably on magnetic tape). Depth intervals will range from 1 to 5 meters.

3. Low-Resolution Data File. Data falling in this category is collected form systems that record on-deck or internally on strip-chart analog records. The data submitted to NODC is digitized from these analog records at variable depth intervals, usually in excess of 5 meters.

Allocation of data to a particular file depends on the resolution of the data collected. Resolution is a function of the interval to which the data is, or can be, digitized and hence is a function of recording method and type of sensor used.

Therefore, it is not always possible to rigidly classify all types of STD data in this manner: e.g., with careful digitization techniques, strip-chart analog records could be digitized at sufficiently close intervals to enable their inclusion in the High-Resolution File.

The accuracy of the data in any of these files will be dependent on the techniques used to collect the data, the quality and quantity of at-sea standardization of data collected, and the type of sensor used.

Two other important files will be maintained by NODC:

4. Expanded Standard-Level Data File. Data in this file may be gleaned by NODC from any of the foregoing files and is compatible with the serial hydrographic data already archived at NODC. The originator may interpolate in-house and submit only the standard-level synopsis of his STD data. The depth interval is similar to NODC's standard levels, except that, due to the continuous nature of the original record, the maximum interval will be no more than 100 meters. This file may well be that most frequently requested by the secondary user and it is essential that it be incorporated into NODC's serial hydrographic cast files. The expanded standard-level data from most STD stations will be of comparable length to the serial-cast data as there will be no regular "observed levels" to file with them. It is important, however, that observed data from bottle samples collected with the original STD station for standardization or other purposes, be contained in the Expanded Standard-Level Data File and indicated as such. Any chemical data obtained this way should also be included.

5. Compressed Data File. It may become impractical for NODC to maintain all of the data submitted for the Ultra-High, High- or Low-Resolution Files. Some method of compression will be necessary in the interests of economy and space. Some originators will already have compressed their data by digitizing analog records to include only inflection points.

Format of STD Data Submitted to NODC Files - It is recommended that:

1. The preferred source of the STD data be that obtained from the descent profile.

2. Representation of STD data submitted to NODC be in character form, either on 7-track BCD (binary coded decimal) or 9-track EBCDC (expanded BCD). This applies particularly to high-resolution data.

3. Digitized analog records or expanded standard-level data interpolated by the originator may be more easily submitted on punched cards in the usual NODC serial-cast format.

4. Units used be degrees centigrade ($^{\circ}$C), parts per thousand (^{o}/oo), meters (m), meters per second (m/sec), milliliters per liter (ml/l), millimho per centimeter (mmho/cm), and microgram-atoms per liter (Ug-at/l) for most chemicals. Pressure units (decidar) may be used rather than depth (meters). If meters are preferred the values used in the conversion of pressure unit must be given.

5. Temperature be reported to 0.001^{o}C, salinity to 0.001^{o}/00. Although the last place may not be compared with other stations' data it would have some internal precision significance.

6. Calculated parameters be sigma-t (σ_t), sound velocity (V), specific volume anomaly (δ), dynamic depth (D), and potential temperature (T_p).

Documentation Submitted by Originator - It is recommended that documentation supplied by the originator, in addition to the usual header information (station location, ship, cruise, time, etc.) to NODC include:

1. Source of data (digital magnetic tape, digitized analog record, etc.).

2. Type of STD instrument; type and number of sensors and serial numbers.

3. Profile submitted (descent or ascent profile).

4. Lowering rate(s).

5. Number of classical standardization points collected.

6. Methods of collecting standardization data (multi-sampler, Nansen-bottles on same/separate cable, etc.).

7. All classical data collected with the particular station by multi-sampler, Nansen-bottles, etc., including chemical data.

8. Corrections applied to the final data.

9. Methods of smoothing, "despiking," filtering.

10. Any unusual circumstances that may have a bearing on the accuracy of the data.

FILTERING AND/OR COMPRESSION METHODS TO BE APPLIED TO STD DATA BY NODC

Filtering - Filtering generally implies a removal of certain parts of the data. In the case of STD data, spikes that may be caused by instrumental "noise" or mismatch of sensor time constants are removed. Included under this heading may also be smoothing and interpolation techniques which many collectors apply to their raw data to produce final data at even depth increments. For filtering it is recommended that:

1. Filtering to remove spikes caused by instrumental "noise" or mismatch of sensor time constants be performed by the originator and not by NODC.

2. Interpolation and smoothing techniques applied to raw data to produce final data at even depth increments be applied by the originator and not by NODC.

NODC has had much experience in interpolating serial hydrographic data using Lagrangian techniques to produce standard levels. In general, STD data submitted to NODC to be placed in the Expanded Standard-Level Data File will have data points close enough to allow a linear interpolation method to be used. However, Lagrangian methods may still be necessary in dealing with digitized analog records whose intervals will not be evenly spaced. Therefore, it is recommended that:

3. Data coming from Ultra-High or High-Resolution Files be linearly interpolated by NODC for inclusion in the Expanded Standard-Level File.

4. NODC apply Lagrangian techniques where necessary in dealing with digitized analog records whose intervals will not be evenly spaced.

Compression Methods - Data compression means the elimination of data points that fall between maxima and minima in the original data record. This method enables NODC to efficiently store large volumes of STD data. The criterion that determines the compression method to be used is that the original record can be reconstructed from the compressed record without introducing errors more significant than the precision of the original data. Therefore, it is recommended that:

1. NODC apply the compression method that was recommended by the NODC/NASCO STD Workshop of September 16-18, 1968, and subsequently the ICES Recommendations.

2. Compressed data be separately filed, as explained in the section on Compressed Data File.

TYPES OF QUALITY-CONTROL PROCEDURES TO BE APPLIED BY NODC AFTER RECEIPT OF STD DATA

For quality control procedures, it is recommended that:

1. Quality-control procedures be applied by the originator of the STD data and not by NODC.

2. NODC develop a quality-control indicator system. This system should be based on information supplied by the originator as listed in the section on Documentation Submitted by Originator; such as, if no classical data points were collected, or if the data points were obtained from bottles on a cast separate from the STD or during a different profile from that being corrected, the STD data will still be archived, but with appropriate quality indicators. These data-points are still of value; although it would be the secondary user's responsibility to judge the suitability of the listing in regard to his own needs.

INFORMATION NEEDED BY SECONDARY USERS OF STD DATA REGARDING DATA COLLECTION METHODS -

It is recommended that the secondary user be supplied by NODC with all of the pertinent documentation submitted by the originator of the STD data. This will allow the secondary user to evaluate the data's suitability for his own needs.