

POSSIBLE ALTERNATIVE SOLUTIONS FOR
ESTABLISHING ICCAT DATA CENTER (S)

By P. M. Miyake
ICCAT Assistant Executive
Secretary

I. BACKGROUND

In 1973, the Secretariat proposed a centralized data bank or data archive system (SCRS/73/7). In that document, the following three alternative plans were proposed:

- 1) Increase the Secretariat staff and funds for establishing a data bank at the headquarters.
- 2) Assign work to working groups (one scientist or institution for each working group for the major species, responsible for data processing.
- 3) Contract an outside organization.

The opinion of the scientists at that time was that "the system currently in effect is satisfactory, whereby each national office completes its data and exchanges them, if necessary". (1973 SCRS Report - Appendix IX)

As the Secretariat commenced the biostatistical tasks in 1976, we encountered some difficulty in the availability of data and in reorganizing some of the data into a desired area/time stratum. Recently, the Convener of the Sub-Committee on Statistics suggested that a new data center system for easy retrieval of the data be considered. After the Assistant Executive Secretary discussed the subject with the Convener of the Sub-Committee, the Secretariat distributed Circular 1976/17, dated August 17, 1976, requesting information on the types of data compiled by each national office. This paper summarizes the responses to that circular, presents possible alternative plans for the establishment of a data center, and discusses advantages and disadvantages of these plans.

II. PRESENT STATUS OF THE DATA STORAGE

The attached Table 1 summarizes the types of data which are presently stored at each national laboratory. The summarization is based on the responses to the request made in Circular 1976/17.

Table 2 summarizes where the data for the major types of fisheries are stored.

A summary of these two tables is as follows:

- 1) According to the information we received, all the major national laboratories have their own data stored. Besides, CNECO/COB (Brest, CRO (Abidjan), CRO (Dakar), FSPRL (Shimizu), FRBC (St. Andrews), NMFS (La Jolla), NMFS (Miami), and the ICCAT Secretariat have computerized at least part of these data.
- 2) It seems that NMFS (La Jolla) has initiated a data bank system in which all the data available in publications and in "Data Records" have been stored, regardless of country, species, etc.
- 3) Also, CNECO/COB (Brest) has compiled in their computer system the C/E data for all long-line fisheries and all albacore biological data available.
- 4) FSPRL (Shimizu) also has computerized all the longline data, regardless of flags.

III. ALTERNATIVE PLANS FOR THE POSSIBLE DATA CENTER(S)

- 1) To establish a data center under the management of the Secretariat.
 - a) Madrid-based
 - i) The Secretariat to be facilitated with a computer.
 - ii) The Secretariat could form a contract with a computer firm and have a terminal in its office.
 - iii) Form a contract with a computer center; all the data and the work could be done at the center, while management could be done by the Secretariat.
 - b) Locate the center at a place other than Madrid.
- 2) Contract another organization under the supervision of the Secretariat
 - a) A commercial firm
 - b) A scientific institute

- 3) Data centers based at various national institutions, while the Secretariat assists in the exchange of information.

IV. ADVANTAGES AND DISADVANTAGES OF THE ALTERNATIVE PLANS

1) Data center managed by the Secretariat

This setup would probably provide the most satisfactory data management and retrieval system. This system would require a considerable amount of work and funds in installing the data and continuously maintaining and updating them. It also requires a lot of cooperation by the national sections in providing all the data to the Secretariat, so that they can be prepared for the data entry.

Within this scheme, there are two alternatives, which are:

- a) Locate the data center in Madrid
- b) Locate the center at another site

To locate the center in Madrid (Plan 1-a) would have an obvious financial advantage of being at the same location as the Secretariat, as well as the advantage of easy data management. The Secretariat had has experience with three different computer firms in Madrid in the past. Unfortunately, the facilities provided by these three firms are not adequate enough to consider establishing a data center with any of them, in terms of i) easy access to the computer, ii) turn-around time, iii) facilities, and iv) accuracy in card punching, etc.

To locate the data center in a place other than Madrid (Plan 1-b) would have certain disadvantages, such as the operation would be more costly and the data management would be difficult (unless the terminal to the computer is placed at the Secretariat). On the other hand, we might be able to find a location where we would have much easier access to the computer with much shorter turn-around time.

A major disadvantage of Plan 1 is that the establishment and maintenance of the data center will be very costly. According to a rough estimation, it would require at least three people (one programmer, one statistician, and one statistical aid) for a one-year period to initiate the system, as well as an easily accessible computer. Later on, we would need one full-time statistician-programmer to maintain the system. This estimate is based on the assumption that the work can be done in Madrid. If the center is located at a place other than Madrid, considerable travel expenses might be required. Naturally, in addition to the above expenses, significant funds must be reserved for the computer time, card punching, printings, mailings, etc.

- ##### 2) Contracts with other organizations
- a) Commercial firm

This would be just a modification of Plan 1. In this case the budget could be even larger than that for Plan 1, even though additional personnel at the Secretariat will not be necessary. A definite disadvantage is that a commercial firm would not necessarily be familiar with the nature of the data, and processing would take more time. In the end, the Secretariat might have to be as much involved as in Plan 1.

- b) Contract a scientific institute

If we take the alternative of Plan 2, this may be more realistic. The scientific institute to be contracted will most likely be one of the tuna research centers at which some data have been installed already, for example, CNEOX/COB (Brest, NMFS (La Jolla), or FSPRL (Shimizu). The preparation of the data entry can be done at the Secretariat or at the institute by the Secretariat staff, or by the institute's own staff. Then, the Commission would provide the funds to process these data. If necessary, the Secretariat might have to provide a programmer or a statistician to do the work. The cost to the Commission would depend largely on how much of the work the institute can absorb in its own budget.

It should be pointed out that this system would give considerable advantage to a particular laboratory to which the work is contracted, because that laboratory will then have much easier and quicker access to the data.

3) Centers located at various institutions

Since data filing systems have already been established in many laboratories, (even though none of them are really complete), it would be quickest and most economical to distribute the responsibility among the national laboratories. However, if this system is adopted, promotion of the exchange of data will be essential. This might be achieved in the following way:

a) The Sub-Committee on Statistics or a new "ad hoc" committee on data management could meet for an extended period in the immediate future to review the available data and agree upon the data input and output formats. These formats should be very detailed and very definite, in terms of species to be entered, gear breakdown, area breakdown, time, etc.

b) Responsibilities among each laboratory should be clearly assigned and agreed upon.

c) Each laboratory would complete its data bank according to the scheme agreed upon under b). Much inter-laboratory cooperation is needed since some raw data have to be exchanged between one laboratory and another to complete its data bank.

d) After the completion of the data banks, each center should then provide summary tables for all the data for which it is responsible.

e) The Secretariat would assemble these tables and publish them. This publication series will be similar to the Data Record, but more complete and, more importantly, all the data will be presented in a uniform and consistent manner. Perhaps a series can be issued for each species so that the scientists do not have to refer to the various volumes to find fragmental information.

f) The system should be maintained and updated every year by the national laboratories. Once the initial publication of all the pertinent data has been completed, only annual, supplemental data would be published each year.

The advantage of this system is obviously that it is the most economical due to the utilization of already-established data banks. One big disadvantage is that all the data are not combined in one system. Therefore, if one scientist wants to work with all the data on Atlantic tuna, he has to enter all the summary tables in his computer. This system could also result in some increase in the budgets of some national laboratories, while the increase in the Commission budget would be minimal.

However, if this plan should be adopted, the biostatistical work or any other statistical tasks assigned to the Secretariat would still experience the considerable drawback of the inavailability of data at the headquarters.

V. SOME CONSIDERATIONS TO BE MADE

There could be many more solutions which fall between the plans suggested here. At any rate, realizing any plan will be very costly, in one way or another, and once a system is adopted, it would be difficult and wasteful to change it. For this reason, a decision on a definite plan should not be made in haste, but with extensive preliminary study.

Here below, a suggestion is made as to a time schedule which might be followed. Naturally, such a schedule could vary greatly, depending on which plan is adopted.

i) 1976 SCRS meeting - Tentative agreement on a plan. Formation of a working group to study the plan.

ii) January-February, 1977 - The working group could meet to study the plan carefully. Formats of data input and output could also be agreed upon.

iii) March-July, 1977 - Each national office could prepare data for an agreed year, according to this format, and send them to the designated site of the data center(s).

iv) August-November, 1977 - Test run these data for one year only. Output should be presented at the next SCRS meeting (1977), so that a final decision could be made at that time.

v) The Commission should take the plan into consideration when estimating the budget for the following biennial period.

TABLE 1. Types of data stored at each laboratory.

| INTERNATIONAL ORGANIZATION | TYPE OF DATA | FLAG | SPECIES | UNIT | GEAR | GENERAL AREA | SMALLEST AREAL UNIT | SMALLEST TIME UNIT | YEARS | APPROX. COVERAGE | HOW STORED | FORMS AVAILABLE | COMPUTER USED | REMARKS (Data Source) |
|-----------------------------------|---------------------------|---|-------------------|---------------------------------------|---------------------|----------------------|-------------------------------|----------------------------|-----------------------------------|------------------------------|-------------------|-------------------|----------------------------|---|
| ICCAT | Noninal catch | All | All | MT | All | Whole Atl. & Med. | 6 regions for Atl. | Year | 1965-75 | 100% | Card & mag. tape | Card & mag. tape | IBM 360/40 | Submitted from national offices |
| | Landing | Kor, Pan, China | YF, Alb, BE | MT, by boat | LL | Whole Atl. | By port | Biweek. | 1973-76 | 50-90% | " | " | -- | Collected at trans-shipment ports by the Secretariat |
| | C/E | Kor, Pan, China | YF, Alb, BE | MT, no. of hooks | LL | Whole Atl. | 5° x 5° | Month | late 74-1976 | 60-90% | " | " | -- | Collected at trans-shipment ports by the Secretariat |
| | Size freq. | Kor, Pan, China | YF, Alb, BE | FL/LDI 1 cm int. | LL | Whole Atl. | 5° x 5° | Month | late 74-1976 | -- | " | " | -- | Collected at trans-shipment ports by the Secretariat |
| CNEO/COB Brest-France | Catches, effort, landings | France | BF, Alb | Numbers, ton | Trol. BB | NE Atl. | 1° x 1° | Week, month | 1967-76 | variable 20-100% | Card & mag. tape | -- | IBM 1130, H.P. + CII 10070 | Fisherman, logbooks, agency (national), professional cooperators |
| | C/E | Japan, part Korea | All | Numbers | LL | Whole Atl. | 5° x 5° | Tri-mentor/month | Jap 1956-1973 China 67-74 (70) | as published by Govt. Agency | " | -- | " | Laboratories, national scientists, ICCAT Secretariat |
| | Size freq. | France, Japan: 56-73 Chian(Taiwan) 73-74 | Alb Alb Alb | FL (TL) " " | Surf. LL " LL | Whole Atl. " " | 1° x 1° 5° x 5° 5° x 5° | Week Mon/Qtr Mon/Qtr | 1967-76 1956-75 1956-75 | 10-20% 5-10% 5-10% | Card " " | -- | -- | Sampling from markets; scientist survey; national agencies; ICCAT |
| CRO-Abidjan ORSTOM-Ivory Coast | C/E | FIS, Morocco | YF, BE, SJ | 100 Kilos | BB, PS | Trop. Atl. | 1° x 1° | 2 Wks. | 1969-75 | Catch-100% C+E-75% | Cards & mag. tape | -- | IBM 370/145-OS/VS | |
| | C/E | Kor-China Japan | YF, BE | Ton | LL | Atl. | 5° x 5° | 2 Wks. | 1973-74 | Sent 80% | Cards | -- | -- | The data are directly sent to ICCAT after 1975 |
| | Measurements | FIS, Morocco | YF | Cm. | BB, PS | Trop. Atl. | 5° x 10° | Month | 1969-75 | 90% | Cards & mag. tape | -- | -- | |
| CRO DI-Senegal | Sampling | All | SJ | -- | All | East Atl. | 5° x 10° | Qtr. | 1969-75 | -- | Cards | -- | IBM 370/45 | All data on C/E for YF, SJ, BE landed at Dakar are centralized in CRO-Abidjan |
| FRB -St. Andrews, Canada | C/E | Canada | SF | No. of fish/day | Harpoon | N.W. Atl. | 1° x 1° | Day | 1958-70 | 30-50% | Cards & mag. tape | Cards & mag. tape | H.P. 3000 | |
| | C/E | Canada | SF | No. of fish & hooks by pos. & date | LL | N.W. Atl. | 1° x 1° | Day | 1962-70 | 50% | Cards & Mag. tape | Cards & Mag. tape | " | |
| | Size Freq. | Canada | SF | Dressed wt(Lbs) & Dressed length (cm) | LL/Harpoon | N.W. Atl. | 1° x 1° | Trip (10-14 days) | 1958-70 | 10% | Cards | Cards | " | |
| | Morphometrics | Canada | SF | Up to 30 measurements/fish (cm) | LL/harpoon | N.W. Atl. | Indiv. position | Day | 1962-70 | -- | " | " | " | 300 fish total in detail, otherwise 5 measurements/fish on 500-1000 fish/yr. |

| INTERNATIONAL ORGANIZATION | TYPE OF DATA | FLAG | SPECIES | UNIT | GEAR | GENERAL AREA | SMALLEST AREAL UNIT | SMALLEST TIME UNIT | YEARS | APPROX. COVERAGE | HOW STORED | FORMS AVAILABLE | COMPUTER USED | REMARKS (Data Source) |
|------------------------------|-------------------------------|-------------------------------|------------------|---|--------|--------------------|--|--------------------|-----------------------|------------------|-----------------------------|-------------------------|----------------|---|
| FRU - Tema, Ghana | Nominal Catch | Jap-Kor Pan-Ghana | YF,SJ,BE, LT, FT | MT | All | E.Trop. Atl. | E.Trop. Atl. | Month | 1972-present | 100% | Manual | Manual | -- | Fisheries Research Unit - Tema, Ghana |
| | Landing | " | " | MT by boat | " | " | " | " | " | " | " | " | -- | Port - Tema, Ghana |
| | C/E | " | " | MT & days | BB | " | " | " | " | " | " | " | -- | " " |
| | Size freq. | " | YF, SJ, BE | FL/cm Inter. | All | " | " | " | " | 43% | " | " | -- | " " |
| FSFRL - Shimizu, Japan | C/E | Japan | All | No. of fish - No. of hooks | LL | Atl. & Med. | 5° x 5° | Month | 1956-74 | est. | Mag. tape | Mag. tape | HITAC 8450 | |
| | " | " | " | MT, no. of fish, days | BB | Atl. | 1° x 1° | " | 1969 1973-75 | 90% < | " | " | FACOM 230 | |
| | " | " | " | MT, no. of hauls | PS | Atl. | 1° x 1° | " | 1967-74 | 100% | " | " | HITAC 8450 | |
| | " | China (Taiwan) | " | No. of fish, no. of hooks | LL | Atl. | 5° x 5° | " | " | est. | " | Mag. tape | " | Data series of Taiwanese catch & effort were re-arranged. |
| | Size freq. | Japan | Tunas | cm, fork length | LL | Whole Atl. Med. | 0F, SJ-5° x 10° Other species- 10° x 20° | Qtr. | 1970- | Vari-able | Mag. tape | Table reported to ICCAT | HITAC 8450 | Alb & SJ 1 cm. intervals Other species 2 cm " |
| FSFRL (CONT'D) | Size freq. | Japan | Bilifishes | cm eye-fork | LL | Whole Atl. Med. | 10° x 20° | Qtr. | 1970- | Vari-able | Mag. tape | Table reported to ICCAT | HITAC 8450 | 5 cm intervals |
| | " | " | Tunas | cm fork length | Surf. | " | 5° x 10° | Month | " | " | " | " | " | SJ- 1cm intervals Other species - 2 cm inter. Data format of MT is not historically uniform |
| I.P. - Santos, S.P. Brazil | Catch & Landing | Brazil | All | MT | LL | South of Brazil | 2 subareas | Month | 1969-76 | 100% | Cards | Cards | -- | Logbook types |
| | C/E | " | YF, Alb, BE, SF | No. wt & no. of hooks | LL | " | " | " | " | " | " | " | -- | " " |
| | Size freq. | " | " | MT | LL | " | " | " | 1969-75 | -- | " | " | -- | Landings |
| | " | " | YF, Alb. | FL | " | " | " | " | 1976 | -- | " | " | -- | 2 cm intervals - landings |
| NMFS-La Jolla, Calif. U.S.A. | Nominal Catches ^{1/} | All | All | MT | All | Entire Atl. & Med. | 20° x 20° | Year | 1960-75 | 100% | Cards Mag. tape (data base) | Cards Mag. tape | UNIVAC 1108 | ICCAT Stat. Bulletin-Vol.5 |
| | C/E | Japan, USA, FIS, China, Korea | All | No. fish, MT hooks, hauls, sets, days fishing | LL, PS | Entire Atl. | 1° x 1° | Month | 1956-75 ^{2/} | 50-100% | " | " | Burroughs 6500 | Fish. Agency, Japan, ICCAT Data Record, Vol 1-7, ICCAT Stat. Series-1; Others |
| | Size freq. ^{3/} | All | YF, BE, Alb, SJ | cm, Kg, ago-Mon, year | All | " | 5° x 5° | " | 1956-75 | ? | " | " | UNIVAC 1108 | ICCAT Data Records, Vols.- 1-7, ORSTOM reports, others |

1/ Data exist on a Data Management Language (DML) data base.

2/ Varies from fleet to fleet.

3/ DML data base in process of creation.

TABLE 2. Summary table on the status of data storage for the major fisheries

| INTERNATIONAL ORGANIZATION | TYPE OF DATA | FLAG | SPECIES | UNIT | GEAR | GENERAL AREA | SMALLEST AREAL UNIT | SMALLEST TIME UNIT | YEARS | APPROX. COVERAGE | HOW STORED | FORMS AVAILABLE | COMPUTER USED | REMARKS (Data Source) |
|----------------------------|---------------------|------|---------|----------------------------|------------|--------------|---------------------|--------------------|---------|------------------|----------------------|---|---------------|----------------------------|
| NMFS - Miami, Fla. U.S.A. | Catch ^{1/} | USA | BF | Kg. | All | N.W. Atl. | 1° x 1° | Day | 1974-75 | 10-90% | Tape | Tape | UNIVAC 1108 | |
| | C/E-sport fish. | " | " | no. per angler, boat, line | Rod & reel | " | " | " | 1975 | 10% | 7 track 800 BPI | Odd parity FL data; even parity BCD avail. | " | |
| | C/E purse seine | " | " | MT | PS | " | " | Hour | 1974-75 | 100% | Tape | Tape | " | |
| | Size freq. | " | " | Cm Int. lb. Int. | All | " | " | Day | 1974-75 | 3-95% | Tape 7 track 800 BPI | Odd parity FL data; Even parity BCD avail. | " | |
| | Morphometrics | " | " | - | All | " | " | " | 1950-75 | - | " | " | " | Large amount of Mor. data. |

^{1/} Data exist on a Data Management Language (DML) data base.

| TYPE OF DATA | AREA | SPECIES | GEAR | FLAG | LABORATORIES | REMARKS |
|---------------|---------------|------------|--------------|-------------------|--------------------------------------|------------------------|
| Nominal catch | All | All | All | All | ICCAT Secretariat NMFS (La Jolla) | 6 regions 20° x 20° |
| C/E | Trop. Atl. | YF, BE, SJ | BB, PS | FIS | CRD (Abidjan) NMFS (La Jolla) | 1° x 1° |
| | " | " | " | American Japan | NMFS (La Jolla) | " |
| | " | " | " | Korea + Pan. | FSFRL (Shimizu) | " |
| | Temp. Atl. | All, BF | Troll + BB | France | CNEXO/COB (Brest) | 1° x 1° |
| | " | " | " | Spain | NMFS (Miami) | 74-75 |
| | " | BF | All | U.S.A. | FRBC (St. Andrews) | 58-70 |
| | " | SF | " | Canada | FSFRL (Shimizu) | 55-73 |
| | Entire Atl. | All | LL | Japan | CNEXO/COB (Brest) | 55-73 |
| | " | " | " | China | NMFS (La Jolla) | 56-73 |
| | " | " | " | " | FSFRL | 67-74 |
| | " | " | " | " | NMFS (La Jolla) | 67-74 |
| | " | " | " | " | Secretariat | 74-76 |
| | " | " | " | Korea + Pan. | CNEXO/COB (Brest) | In part |
| | " | " | " | " | CRD (Abidjan) | In part |
| | " | " | " | " | NMFS (La Jolla) | In part |
| | " | " | " | " | Secretariat | 74-76 |
| Size freq. | Trop. Atl. | YF, BE, SJ | BB, PS | FIS | CRD (Abidjan) | 5° x 10° |
| | " | SJ | " | FIS | NMFS (La Jolla) | 5° x 5° |
| | " | YF, BE, SJ | BB, PS | FIS | CRD (Bakar) | 5° x 10° |
| | " | YF, BE, SJ | BB, PS | American | NMFS (La Jolla) | 5° x 5° |
| | " | YF, BE, SJ | BB, PS | Japan | FSFRL (Shimizu) | 5° x 10° |
| | " | " | " | Spain | " | 70-75 |
| | " | " | " | Korea + Pan. | NMFS (La Jolla) | ? |
| | Temp. Atl. | All, BF | Troll + BB | France | CNEXO/COB (Brest) | 1° x 1° |
| | " | " | " | U.S.A. | NMFS (La Jolla) | 5° x 5° |
| | " | BF | All | " | NMFS (Miami) | 74-75 |
| | " | All, BF | Troll + BB | Spain | " | ? |
| Entire Atl. | All | All | LL | Japan | FSFRL (Shimizu) | 5° x 10° (BF, SJ) |
| | (Alb) | " | " | Japan | CNEXO/COB (Brest) | 10° x 20° (Others) |
| | Alb. | " | LL | " | NMFS (La Jolla) | 5° x 5° |
| | Alb + YF + BE | " | " | China | CNEXO/COB (Brest) | 5° x 5° |
| | " | " | LL | " | Secretariat | 1° x 1° |
| | " | " | LL | Korea | Secretariat | 74-76 |
| | SF | " | LL + Harpoon | Canada | FRBC (St. Andrews) | 74-76 |
| | " | " | " | " | " | 58-70 |