ASSESSMENT OF POTENTIAL FOR IMPROVING AGRICULTURAL DATA COLLECTION AND PROCESSING

By

Sandra Rowland
General Surveys Branch

and

William Stuart
Computer Applications Branch
International Statistical Programs Center
Bureau of the Census

February, 1981

INDONESIA

Assessment of Potential for Improving Agricultural Data Collection and Processing

I. Scope of the Report

A. Introduction and Purpose 1

B. Summary of Findings and Recommendations 3

II. Review of Present System of Agricultural Statistics 10

Introduction 10

A. Central Bureau of Statistics 11

1. Food Crops 11

a. Area 14

b. Yield and Production Forecasts 14

2. Estate Crops 15

3. The Future of Processing at CBS 15

4. Pilot Projects and Other Activities of CBS Related to Agricultural Statistics 15

a. Smallholder Estate Crops Pilot Surveys 15

b. Livestock Survey 16

c. Farm Household Survey 16

d. Price Statistics 17

e. Food Consumption 17

f. Remote Sensing/Area Frame Construction 17

B. Department of Agriculture 17

1. Directorate General of Estate Crops 18

2. Directorate General of Fisheries 20

3. Directorate General of Forestry 22

4. Directorate General of Animal Husbandry 23

5. BIMAS 23

6. Agency for Agricultural Research Development (AARD) 24

III. Assessment of Institutional Capability to Modify the Present System of Agricultural Statistics 25

A. Improvement of the Quality and Scope of the Present System of Agricultural Statistics 25

1. Major Complaints 25

2. Methodological Problems and Recommendations 26

a. Scope of Present Agricultural Statistical System 26

b. Frequency of Collection 27

c. Statistical Methodology (Sampling vs Census) 27

d. Quality Control in the Field 29

e. Data Processing 30

3. Practical Constraints 30

a. Institutional Requirements 30

b. Resource Constraints 32

B. Improvement of DOA's Ability to Collect and Process Data 32

1. Rationale 32

2. Types of Training 33

a. Sampling 33

b. Survey Statistics 33

c. Data Processing 33

d. Short-term, Long-term and On-the-job Training 34

3. Requirements for Successful Training and Absorption of Trained Personnel 34

a. Central Control and Design of a Purposive Training Scheme 34

b. Calendar Time and Organization of Training 34

4. Suggested Training Program 35

5. Description of On-the-job Training 36

a. Sampling 36

b. Survey Statistics 36

c. Data Processing 37

C. Improvement of DOA's Ability to Process Primary Data 38

1. Establishment of a Computer Center within AARD 38

a. Hardware 38

b. Software 41

c. Staff Training 44

d. The Computer Facility 46

e. Purpose of Computer Facility 48

PERSONS CONTACTED 49

LIST OF TABLES AND FIGURES

Table 1. Central Bureau of Statistics Data Collection Activities 12

Table 2. Estimated Keying and Storage Requirements for Area and Production Statistics 31

Table 3. Financial Requirements for Suggested Training Programs for Collection and Processing of Primary Data in DOA (in U.S. $) 36

Table 4. Estimated Keying and Storage Requirements for BIMAS Data 47

Figure 1. Organization Chart - Department (Ministry) of Agriculture 19

Figure 2. Flowchart of Data Processing System 40

-1-

INDONESIA

ASSESSMENT OF POTENTIAL FOR IMPROVING AGRICULTURAL DATA COLLECTION AND PROCESSING

I. SCOPE OF THE REPORT

A. Introduction and Purpose

In 1977 a five-year Agricultural Planning and Administration Project was initiated by AID and GOI. The executing agency is the Department of Agriculture (DOA) and the Project Manager for the GOI is Chief of the Bureau of Planning of DOA. The purpose of the project is to "upgrade the planning and programming capability of the DOA by helping to institutionalize within the DOA the ability to carry out effective Agricultural Policy Analysis and the capacity to design and evaluate appropriate developmental program and project alternatives". 1/ Although the Project Paper is decidedly vague with respect to the specific activities which should be undertaken to accomplish the purpose, it does emphasize data analysis and planning as opposed to primary data collection and processing. Any references to data concentrate on their management rather than their production. Specifically the improvement of planning is to be accomplished by "(a) technical assistance for planning, data programming and training under a host-country contract between a U.S. University and DOA and (b) long and short-term academic training of some 380 DOA planners ... Improving the data management system will be accomplished by (a) centralizing the data processing functions in one agency of DOA, (b) constructing a new building in Jakarta to house this activity (c) installing the needed data processing equipment and (d) improving the data processing functions through technical assistance." 2/ No concern for the improvement of agricultural statistics in terms of their scope or quality is evident in the project paper, however, when the contract was signed between a U.S. University and the DOA a strong component for the improvement of agricultural statistics was built into the job description of one of the resident advisors.

The project is more than two years into its life and it has focused on training economic analysts and planners, providing a building for a computer, obtaining software for the computer and training programmers for the computer center which is being centralized in the Agency for Agricultural Research and Development (AARD) of the DOA. To date little attention has been given to the collection and processing of primary agricultural data.

The U.S. Census Bureau (BuCen) under its central contract to DS/DIU/DSB of AID has been requested to evaluate the possibility of building a primary data collection and processing component into the project based

---

1/ Indonesian Agricultural Development Planning and Administration, Project Paper, Dept. of State, Agency for International Development, Wash. D.C. Sept. 1977, P. 14

2/ Ibid. pp. 5-6

upon its assessment of the potential for improving agricultural data collection and processing in the GOI. Specifically, in response to USAID's request for BuCen to (1) identify areas of greatest need concerning standardization/computerization of Indonesian agricultural statistics and (2) analyze the feasibility of installing BuCen software, BuCen recommended and USAID agreed to the following scope of work:

(1) Conduct survey of GOI Institutions to determine nature of statistical activities in area of agriculture carried out by these GOI institutions.

(2) Determine quality and reliability of these statistics.

(3) Assess the institutional capacity of the GOI institutions responsible for agricultural statistics to carry out their tasks and improve the reliability of their statistical output.

(4) Conduct an analysis of the principal gaps in the system of agricultural statistics from the point of view of the flow of information required to support planning efforts for the development of the agricultural sector.

(5) Evaluate the possibility of establishing a national system of agricultural statistics.

(6) Recommend a realistic time table and resources required to produce these statistics taking into account the institutional framework in which they will be produced and the technical capability of the institutions involved.

(7) Recommend mechanisms to coordinate activities in the field of agricultural statistics so that the adoption of common standards, classification and nomenclature, evaluation and analysis, become possible.

(8) Assess the current data processing situation in terms of hardware, software, and personnel resources; recommend appropriate software packages; and make plans for providing software and training in its use and future technical assistance in other aspects of data processing.

The present system of agricultural statistics implemented by the Central Bureau of Statistics (CBS) and the Department of Agriculture (DOA) will be reviewed in Section II of this paper in order to provide a background for the recommendations made. An assessment of the institutional capability of the CBS and DOA to modify and/or improve the present system will be made in Section III along with recommendations for modifications and/or improvements.

All observations and recommendations concerning (1) modification or improvement in the system, (2) realistic time tables and resources required (3) coordination of activities and adoption of common standards and (4) the current data processing situation in terms of hardware, software and personnel will be more meaningful if fit into an intelligible framework. Therefore, Section III of this paper is divided into three categories which vary considerably in terms of the scope of intervention and its cost but are not necessarily mutually exclusive. Part A of Section III contains a discussion of what would be required to improve the scope and quality of the agricultural statistics presently collected by CBS and MOA and probably falls out of the purview of the Agricultural Planning and Administration Project. Part B contains a discussion of the requirements for improving the ability of DOA personnel to collect and process agricultural statistics which may or may not be within the purview of the project depending upon the choices made by AID and DOA as to which course the project should take. Part C contains a discussion of the ability of DOA to process and/or manage the data generated by CBS and MOA which is much akin to the original purposes defined in the project paper and the course the project has followed to date.

It must be emphasized that the expertise of BuCen is in the collection and processing of primary statistical data and not in its use. Most of the observations and recommendations made, therefore, are confined to that area of expertise.

B. Summary of Findings and Recommendations

In our view of the quality and scope of the present system of agricultural statistics we found the following:

1) Agricultural statistics are collected by the Central Bureau of Statistics (CBS) which is legally responsible. CBS works in collaboration with the Department of Agriculture (DOA) which shares responsibility for the actual data collection in the subdistricts. The statistics are manually processed by both CBS and DOA in parallel. This causes disparity in their independent results and a duplication of efforts.

2) The scope of the agricultural statistical system is comprehensive in terms of what is generally considered to be an adequate system of agricultural statistics. However, it is extremely difficult (1) to understand how CBS, DOA and BIMAS 1/ interact to produce the body of data available, (2) to determine exactly what variables are collected for what types of crops and livestock and at what frequency, (3) to determine how much duplication of effort is taking place and (4) to determine what data is available for users and in what form (tables, tapes, etc.).

3) The major constraint to the production of accurate, timely statistics is the lack of a sampling frame which precludes the use of sampling for most of the agricultural statistics collected.

1/ For an explanation of BIMAS refer to page 23.

4) The accuracy and timeliness of the statistics collected are not optimal because the scope of work of the field data collectors (mantri) is burdensome given their number and status.

5) Processing of agricultural statistics in CBS is labor intensive and apparently entirely adequate to meet the current needs of CBS. Conversion to a system of computer processing would require a minimum of twenty person-years to accomplish the keying of the data being received annually and several person-years of highly skilled (and scarce) native programmers.

If the Government of Indonesia is serious about improving the quality and scope of agricultural statistics, the following is recommended:

1) A comprehensive inventory of the data being collected is needed. Such an inventory would be useful for cutting back on duplicative efforts and organizing the data available for research and planning. We therefore, recommend that CBS be contracted to develop an inventory of agricultural statistics, because, on behalf of the Government, it acts as a coordinating body of all statistical activities of the other government agencies. We also believe that it is within reason to suggest that the effort be paid for out of the loan funds from the Agricultural Planning and Administration Project because one of the major purposes of the Project is to improve the data management system and any such effort will require an inventory as a foundation for its development.

2) A pilot project for the construction of area frames is already underway at CBS. If the results of the pilot surveys based upon the frames are good, consideration could be given to developing the frame nationwide or at least for the most important agricultural areas. Alternatively, the population census, if it listed variables which determine which households are farm households, could provide a list frame. The upcoming agricultural census, if carried out with full coverage on all islands, could also provide a frame.

3) The chief of the Agricultural Statistics Division of CBS is proposing that a survey be made of the problems of the field data collectors (mantri) at the subdistrict level to find out what the problems are. We support that recommendation and suggest that the survey be directed at finding out the number of mantri available by subdistrict, their workload, the return rate of forms, the problems they encounter in filling out the forms and the suggestions they have for improving the timeliness of the reporting procedures.

4) From what we have been able to learn, the only ways to improve the quality of field work for the eyewitness system are to increase the number of mantri in areas where they are scarce, raise their incentive pay and status, remove from their scope of work matters not related to statistics (In other words, make them all mantri statistik), establish a national priority as to which forms should be returned first, pay them with central funds, cut out duplications of effort, increase the quality and amount of training and supply them with pencils, paper and hand calculators.

5) It seems inadvisable to us that a conversion to computer processing be undertaken before serious thought is given to the overall data collection and processing effort or before a manifest need for more complex processing methods urges the introduction of computer processing.

6) In order to make a significant improvement in the quality and timeliness of agricultural statistics it will be necessary to establish priorities and make available the funds and personnel required to implement sought after changes. The Committee on Agricultural Statistics of BAPPENAS[1] should be responsible for hearing and analyzing criticisms and suggestions related to the scope, quality and timeliness of agricultural statistics. It should establish priorities for their improvement; for example, it should choose between improving area statistics or increasing the data generated on high-value crops. (From what we could gather it's already performing this function.) A national mandate is absolutely necessary for making an impact upon the system of agricultural statistics.

7) We also recommend that rather than looking for ways to increase the scope of agricultural statistics, efforts be made to set priorities for the improvement of the accuracy and timeliness of the data which are already being collected. Any requests to increase the scope of statistics should be evaluated in terms of the system as a whole (within the Agricultural Statistics Committee) so that trade-offs among scope, quality and timeliness are considered and the necessary funds and personnel for changing the system are made available.

8) The Committee should also be responsible for coordinating the relations between CBS and DOA so that redundancy of work is not permitted. For example, both CBS and DOA are considering the development of a computerized agricultural data base.

9) Any program for the improvement of agricultural statistics should be implemented through CBS with the cooperation of DOA because CBS has a considerable amount of experience in sample survey and census design and implementation and automatic data processing. Personnel in CBS are aware of the problems with the present system and on many fronts are anticipating possible means of improvement through pilots of sample surveys, area frame pilots, a survey of the problems of the mantri, coverage of the upcoming agricultural census and its use or the use of the population census as a sampling frame.

In our evaluation of the ability of the DOA to collect and process primary statistics we found the following:

1) Although the DOA is greatly responsible for collection of agricultural data in the field, it has little or no experience in sampling, survey and census design and implementation and automatic data processing. It should take a secondary role in the area of agricultural statistics for the time being.

---

[1] BAPPENAS is the Indonesian acronym for the National Planning Agency.

2) The five Directorates General (DGs) and the Agency for Agricultural Research and Development (AARD) are the most important entities in the DOA which collect primary and secondary data concerning agricultural research and production. They are located in separate buildings throughout Jakarta and the surrounding areas and function autonomously with respect to their data collection and processing activities.

3) With formal and on-the-job training, it might be possible for DOA to improve some of the statistical programs under its control in the areas of forestry, fisheries and livestock; transfer some of the functions presently carried out by CBS to DOA; and/or design and implement special one-time surveys which will fulfill special ad hoc requirements for project and program planning within DOA.

In order to design an intelligent training program which will meet the needs of DOA in the areas of primary data collection and processing it will be necessary for the Agricultural Information Advisory Committee of DOA to determine:

(1) the number of people to be trained in each area
(2) the distribution of people to be trained among DGs and AARD
(3) the amount of time each trainer can be allowed to train without disrupting the on-going work of the DG or AARD
(4) the amount of financial resources they want to commit to each type of training
(5) which potential primary data collection activities (being carried out or planned for implementation) could be used as vehicles for on-the-job training.

We recommend the following basic training program.

1) Seven people, one from each of the five DGs, one from BIMAS and one from AARD, should be trained in agriculture survey statistics for a period of one year at the International Statistical Programs Center (ISPC) of the Census Bureau in Washington, D. C., two people from AARD should be trained in sampling at ISPC or any university chosen by the Committee; and two programmers from AARD should be trained at ISPC for a period of one year in programming and statistical methods.

2) The six[1] people trained in agriculture survey statistics would serve as core resource persons in the area of primary data collection for their respective DGs and as liaisons between the DGs and the samplers and data processors which should be located in the Data Processing and Statistics Division of AARD. The samplers would assist the DGs in the design of samples for their survey activities and the programmers would assist the DGs in development of software systems by programming themselves or by assisting programmers provided by the DGs. This entire group of ten[1] people would provide the core of experts which could form the basis for improving primary data collection and analysis in DOA.

[1] Does not include BIMAS personnel.

3) Those people who have acquired one to two years of long-term training should proceed with on-the-job training upon their return to work and could be assisted by those who have received short-term training. On-the-job training should ideally include technical assistance by CBS advisors or foreign experts so that problems encountered in the initial work undertaken can be smoothly resolved.

4) It will be necessary to specify activities for use as training vehicles in DOA which require and would benefit from expertise in the areas of sampling, survey statistics and statistical data processing. A number of such activities have been tentatively identified in the report for training purposes.

5) In conjunction with experts at CBS or foreign experts, the core of trained people could develop a series of two to four week courses which can be taught to personnel in the planning bureaus of the DGs in relevant areas like questionnaire and forms design, development of training manuals for fieldwork (interviewers' and supervisors' manuals), writing specifications for computer editing and tabulations, editing and computation procedures, systems design for processing of primary data, quality control programs for survey and census implementation, development of training sessions for mantri, etc.

If the DOA and USAID decide that they prefer to concentrate their efforts in improving the ability of DOA to process and/or manage the data already generated by CBS and DOA (without regard to its improvement), we would like to make the following observations:

1) As one researcher we interviewed expressed it, sometimes supply must be relied upon to create a demand. It was surely the hope of those who first envisioned the establishment of a computer facility within the DOA that the availability of such a center would lead to increased use of statistical data in agricultural planning, more use of modeling and linear programming and gradually more sophisticated analysis. It is unusual to see the establishment of such a facility before there are clearly identified users but, in the long term, the realization of such a goal does not appear unreasonable. Undoubtedly, the process of finding and training users of the system will be a gradual and sometimes frustratingly slow process.

2) Although the computer center was designed to improve communications among the DGs in the DOA and serve the entire organization, it is located organizationally at a low level (within the AARD) of DOA and may not have the ability to coordinate and serve the data processing needs of the DOA without a more formal mandate which requires cooperation and use of the facility by the relatively autonomous DSs.

3) There is a scarcity of both computer programming experience and systems analysis expertise in the DOA. The DGs are the most important prospective users of the center but they do not have trained programmers to use it.

4) Unfortunately, the track record for software contracts to design data base management systems is a disappointing one. The more diffused the focus of the project, the more complex is the design and the greater are the demands placed upon the ultimate user of the system to guide the software contractor in the design. If the data base design does undertake to include provision for routinely collected data on crop area, crop yield or crop production, the software contractor is going to have to face directly the issue of what the source of these data will be. As we described in the report, there are two quite distinct and independent sets of edited data at the level of greatest resolution, the subdistrict level. Should the data be taken from DOA sources, there will be the obvious disadvantage that the data subsequently retrieved from the system will differ from the officially published statistics. If the CBS should be chosen as the source for the data perhaps it would be best if such an undertaking were simply postponed until the CBS could be involved in it as well.

The following recommendations may be of use to the DOA and USAID in the establishment and use of the new computer facility in the DOA.

1) The decision to design a data base system for routinely collected data on crop area, production and yield ought to be one reviewed by the Steering Committee for Agricultural Statistics of BAPPENAS prior to any attempt to design such a data base. The success of this software contract, (as with all others) will be largely dependent on the various DOA officials to state, clearly and in considerable detail, the needs which they will bring to the data base system once it is complete.

2) As a need develops, it may be worthwhile for the center to request the installation of COCENTS, a COBOL coded version of the census Tabulation System. Although COCENTS has not been tested on the Honeywell, Level 6 system, the modifications for its successful adaptation are unlikely to be extensive. Once COCENTS has been installed, training for staff of the center and others who will use the package could be arranged either through the CBS in Jakarta or the International Statistical Programs Center.

3) The most serious hardship that this staff will have to contend with is the lack of experienced systems analysts. There is going to be a need for technical assistance in this staff for a number of years, until some members of the staff gain enough experience to take responsibility for the design and direction of programming efforts. One full-time systems analyst is going to be required to organize and manage the work of the center. If programming work, other than that directed by the software contractor (for the design of an agricultural data base), is going to go on there will be a need for at least one other part-time systems analyst who can plan for and direct the available programmers.

4) In our interviews, we have identified two possible training projects that could be undertaken by the four available applications programmers late in 1981 or early in 1982 if funds could be found for technical assistance in directing the work of these programmers. The projects are discussed in the report.

5) Since full utilization of the Computer Center depends upon the availability of trained computer personnel within the various Directorates General of DOA, it would be advisable for the Computer Center to provide in-house training opportunities. The need for trainees to master English prior to attending computer courses abroad is one of the major obstacles to such training. The addition to the Computer Center staff of an experienced senior programmer whose sole responsibility would be the organization and teaching (in Indonesian) of computer classes and workshops would eliminate the need to train prospective programmers in English.

6) The ultimate installation of SPSS is probably going to be critical to the success of the center and every effort should be made to persuade the hardware supplier to put pressure on the contractor responsible for the SPSS conversion. Should SPSS prove ultimately to be unavailable, some alternative package such as P-Stat, which is written in FORTRAN, might prove sufficiently portable to be of use.

7) It might also be advisable for the center to consider the purchase of another compiler or interpreter to provide one of the newer and simpler programming languages that are now available. Both BASIC and PASCAL were designed with the training of new programmers in mind and both would make the system more approachable by individual users such as researchers in AARD. Of these two possibilities, PASCAL would be the better choice because of its more readable format and its greater potential for generalized use.

8) It seems likely to us that the proposed computer center staff with the technical assistance of at least one part-time systems analyst responsible solely for programming design will be adequate to meet the reasonable programming needs of the DOA for several years to come. Given the scarcity of computer programming experience, it is best that the few people with experience be centrally located where they can benefit from the experience of other programmers and from a broad exposure to user needs. Initially, the different Directorates General should be allowed to contract the center's programmers' time. Gradually, as the Directorates Generals' programmer requirements grow, one or more may recognize the desirability of establishing its own data processing staff.

9) There are two potential problems that will face the DOA in the development of a computer programming staff. The first of these is the lack of appropriate computer job descriptions and titles. It will be unsatisfactory after a point to have to fit a programmer or a systems analyst into a statistician's slot. New positions will be needed for keyers, keying supervisors, computer operators, computer programmers and systems analysts. The second of these potential problems is the lure of private computer contractors who are able to offer higher salaries than civil service requirements will allow the government to offer. One approach to this problem is to find a way to offer more competitive salaries. Some governments, facing a similar situation, have found it possible to establish an independent authority able to pay wages exceeding normal civil service wages and bringing salaries into a range that will reduce the attrition to an acceptable level.

II. Review of Present System of Agricultural Statistics

According to Indonesian law, the Central Bureau of Statistics (CBS) is responsible for the collection of basic statistical data. Agricultural statistics are collected by CBS in close collaboration with the Department of Agriculture (DOA) which shares the responsibility for the actual data collection in the subdistrict (kecamatan) for food crops and estate crops, and carries out most of the data collection and processing for forestry, fisheries and livestock.1/

The majority of the agricultural statistics produced in Indonesia by the CBS and the DOA have been, and currently are, manually processed. This processing consists primarily of editing, aggregation and tabular presentation of the aggregated data.

The CBS and the DOA have cooperated in the collection of agricultural data for two decades although the two groups each process this data independently. The disparity of their independent results became a sufficient cause for concern in the 1960's to prompt an order by the Agriculture Statistics Committee within BAPPENAS (the national development planning bureau in Indonesia) designating the CBS as the official government source of statistics on crop area and production. Despite this action, these data continue to be processed in parallel by the two organizations.

Due to the complexity of the relationship, an effort is made in this section of the paper to describe the scope of work encompassed in the agricultural statistical program of Indonesia and the manner in which CBS and MOA carry out the work. Given that CBS is responsible for collecting, processing and publishing the bulk of the statistics produced, its work will be reviewed first.

A. Central Bureau of Statistics

Most of the work on agricultural statistics in CBS is carried out in the Agricultural Statistics Division which employs approximately 85 people. The most demanding group of agricultural statistics is food crops which occupies 50 people full time. Work on estate crops requires 18 people and the rest of the staff works on fishery and livestock statistics.

Table 1 contains a summary of the agricultural data collection activities in CBS. For each type of data collected, the following information is listed: frequency of collection, number of forms, frequency of processing, and publication, geographic coverage, statistical method and data collection method. The following is a brief description of the data collection activities by type.

1/ Agricultural Statistics Development Plan in Indonesia 1979-1984 by Kartono Hardjopertomo, Unpublished mimeograph submitted to Seventh Session of the FAO Asia and Far East Commission on Agricultural Statistics in Bangkok, Thailand, August 1978. p.1.

1. Food Crops

a) Area

Statistics related to land use and area in crops (at various stages of growth) are collected using seven different forms:

SP1A rice paddy
SP1B other staples
SP2 vegetables
SP3 fruits
SP4 area with pests and disease
SP5A actual and potential land use for rice
SP5B tools and machinery

The data are collected nationwide and are based on eye witness accounts of local village chiefs or other relevant local officials. Specifically, the country is organized into 27 provinces which have a number of regencies (kabupaten) which are divided into subdistricts (kecamatan). There are approximately 3200 subdistricts and each has personnel from CBS (mantri statistik) and the Directorate General of Food Crops of DOA (mantri pertanian or tani) who are responsible for recording the eyewitness accounts of local representatives from the villages in their subdistrict.1/ The mantri tani are paid 7500 rupiah per month, above their regular salary of 70,000, to record the information. Mantri statistik are paid directly by the central office of CBS while mantri tani are paid by the provincial offices of MOA.

The statistical method used could be defined as a census because it covers all of the land area in the country without the use of sampling. All of the forms are filled out monthly by the mantri with the exception of SP3 (fruit) which is filled out quarterly and SP5A (land use) and SP5B (capital goods) which are filled out annually. The scope of work and quantity of data collected are both enormous. (Refer to columns 1 through 3 of Table 1 and Table 2).

The field work is divided evenly between the mantri statistik and the mantri tani. As well as filling out the forms specified above, the mantri tani are responsible for filling out other forms, both statistical and administrative, for the DOA. In addition to all of the data collected by DOA and CBS staff, more data specifically related to rice intensification programs are collected for BIMAS (mass guidance in rice cultivations).2/

1/ Data Systems to Support Agricultural Policy Research in Indonesia by Sampe Tonapa, Unpublished Manuscript presented at the International Seminar for the Asian Region on Agriculture Sector Analysis, Nov. 1976. p.3.

2/ See Section II E for a description of the BIMAS program.

TABLE 1: CENTRAL BURE[illegible] STATISTICS DATA COLLECTION ACTIVITIES

TYPE	FREQUENCY OF COLLECTION	NUMBER OF FORMS PER ROUND	FREQUENCY		STATISTICAL METHOD	GEOGRAPHIC COVERAGE	DATA COLLECTION METHOD
			PROCESSING	PUBLICATION			
1. FOOD CROPS							
a. AREA STATISTICS							
SP1A rice paddy	monthly	3200	continuous	annual	census	nationwide	Eyewitness estimates of village chiefs and other local officials recorded by field staff of the branch offices of CBS and MOA (mantri statistik and mantri tani respectively)
SP1B other staples	monthly	3200	continuous	annual	census	nationwide	
SP2 vegetables	monthly	3200	continuous	annual	census	nationwide	
SP3 fruit	quarterly	3200	continuous	annual	census	nationwide	
SP4 pests and disease	monthly	3200	continuous	annual	census	nationwide	
SP5A land use	annually	3200	continuous	annual	census	nationwide	
SP5B tools and machinery use	annually	3200	continuous	annual	census	nationwide	
b. YIELD STATISTICS							
DAFTAR1 List of plots	annually	-	-	-	census	nationwide	Objective crop-cutting measurements carried out by mantri statistik and mantri tani
DAFTAR2 Crop cutting	3 times pe[illegible]	-	-	-	random sample	nationwide	
- rice	yr.	26,666	continuous	annual		nationwide	
- other food crops		10,000	continuous	annual			
2. ESTATE CROPS							
large holders > 25 ha.	annually	1200	annually	annual	census	nationwide	Mail-out questionnaires filled out by managers and returned to CBS
	monthly	1200	continuous	annual	census	nationwide	
3. PILOT PROJECTS							
a. Smallholder estate crops	once	-	once	once	sample survey		Direct Interview of Farmer
(-rubber)	(1976)	-	-	-	-	2 provinces	
(-others)	(1977-78)	-	-	-	-	5 provinces	
b. Livestock Survey	once 1980	60,000	once	once	sample survey	nationwide	Direct Interview of Farmer

TYPE	FREQUENCY OF COLLECTION	NUMBER OF FORMS PER ROUND	FREQUENCY		STATISTICAL METHOD	GEOGRAPHIC COVERAGE	DATA COLLECTION METHOD
			PROCESSING	PUBLICATION			
c. Farm Household Survey	annually	60,000	annually	annual	sample survey	nationwide	Direct Interview of Farmer
d. Price Statistics	?				sample survey	nationwide	Farm level prices obtained through direct interview of farmer
e. Food Consumption (SUSENAS)	annually		annually	annual	sample survey	nationwide	Direct interview of farm household members
f. Area Frame Pilot	1980 1981		once	once	sample survey	2 areas	Direct Interview of Farmer

Data on crop area comprise the largest single mass of primary data collected in the system of agricultural statistics. The forms used for these data were designed jointly by the CBS and the DOA with manual processing of the data in mind. Each is basically tabular in form with column numbers printed to aid in the transcription of the data. One form used to collect crop cutting data (Daftar II) has been designed for computer use. Boxes appear in selected items for the recording of digits, though these boxes are unnumbered.

From the time of its recording at the subdistrict level, data on crop area is distributed to CBS, DOA and BIMAS officials at the regency level. A copy of the recorded data is also sent directly to CBS in Jakarta by mail.

(1) Processing of Area Data in CBS

Data from the subdistrict level is received by mail at the central CBS office in Jakarta. There, a staff of between 50 and 60 people examine the reports, comparing the current response with responses for the corresponding period in past years. Any necessary editing of the data is done in the central office. The data are then aggregated manually at the regency, province and national levels. The work is done on a continual flow basis. Monthly reports are frequently delayed and are often not received at all. (The return rate is approximately 60%, 80% of these from Java.) As a consequence, the final publication of annual data is delayed by a matter of years, the most recent published statistics being 1977. In the interim, data are available for inspection and use by CBS or DOA officials and are distributed in stencil form, when necessary.

(2) Processing of Area Data in the Directorate General of Food Crops

Data destined for the office of the DG of Food Crops in Jakarta passes through DOA offices at the regency and province levels where it is examined and edited by DOA officials at each level. Data passing through the DOA is consequently edited in approximately 350 different DOA offices around Indonesia before it is received at central offices in Jakarta for final aggregation. We were unable to determine either the approximate number of people engaged in this process of editing and aggregation or whether there are written procedures describing this work, due to the failure of representatives of the DG of Food Crops to meet with us at either of two scheduled appointments.

DOA officials believe that the experience of DOA personnel at the regency level is essential to successful review and editing of the data reported at the subdistrict level. Personnel at the regency level are said to be better trained and, being familiar with the region they serve, and better able to determine inappropriate or invalid responses. The DOA official at the regency level also has the advantage that he is in a position to put pressure on the mantri pertanian for completion or correction of the reporting forms. Obviously, there are also disadvantages in this role of

regional officials, since the opportunity arises for adjustment of the reported data by those whose performance may be evaluated, at least partly, in terms of the information being passed on to the DOA at the higher province and/or national levels.

We must assume that the data thus aggregated by the DG of Food Crops is used for the production of routine and ad hoc in-house reports only, since they are not permitted to publish these results as official government statistics.

(3) Processing of the Area Data in BIMAS

BIMAS offices in Jakarta receive edited and aggregated data on crop hectarage from DOA offices at the province level. A staff of five persons in Jakarta is adequate for the production of routine reports though the processing staff is sometimes hard pressed by demands for special reports from the staff of the DOA.

b) Yield and Production Forecasts

In order to obtain food crop yields, plots of 2½ x 2½ square meters are selected randomly and objectively measured through crop cutting by mantri statistik and mantri pertanian. They are paid 750 rupiah per plot in addition to their regular salary. Every year 80,000 rice plots (one for every 100 hectares) and 30,000 plots of other food crops (one for every 200 hectares) are cut. The work is carried out continuously throughout the year and the results are published annually by CBS. Three times a year (February, June and October) the results from the crop cutting survey and the area estimates are used to forecast rice paddy production. The forecasts are calculated manually by CBS using regression.

2. Estate Crops

The work required to collect data on estate (commercial) crops is divided between CBS (mantri statistik) and the Directorate General of Estate Crops of DOA (mantri kebun) according to farm (enterprise) size. Largeholders, over 25 hectares in size, are covered by CBS through mail-out questionnaires. Approximately 1200 largeholders receive a one-page questionnaire each month and an eight-page questionnaire annually. They are filled out by farm managers and returned to CBS. Data from smallholders are collected by the DG of Estate Crops in the same manner that data on food crops are collected (through eye witness accounts from local officials recorded by the mantri).1/ Forms on estate crop production are received by mail at CBS in Jakarta directly from the 1200 estates. A staff of 18 persons is responsible for manually aggregating this data at CBS.

1/ Refer to Section IIB 1 for more detail on the smallholder census carried out by the DG of Estate Crops.

3. The Future of Processing at CBS

The size of the processing staff for agricultural statistics at CBS appears to be adequate for the task given to it and even using current manual processing methods, the staff is able to complete its work accurately in a timely manner. The cause of delay in the publication of statistics is not in processing but in reporting. CBS has planned a survey of its subdistrict personnel to determine the cause for this delay so that informed remedial action may be taken.

Officials at CBS have given thought to the computerization of their agricultural statistics division but their experience in this area urges them to proceed cautiously. They find that unless there is a demand for involved cross tabulations requiring the use of a computer, the process is more efficiently done by manual methods.

Nevertheless, CBS does plan to give serious consideration to future computerization of agricultural statistics but it intends to consider the entire system of data collection and processing before undertaking the task, i.e., DOA officials and mantri will work with data processing staff in the possible redesign of reporting forms once a clear understanding of the nature of the information desired from the field is reached. CBS intends to begin this process in one to two years and will proceed to address each of the current data collection efforts in turn. This conversion process is not expected to be a particularly large effort initially and probably will involve only one systems analyst, one subject matter statistician and several programmers. A computerized agricultural data base may also be developed at that time.

4. Pilot Projects and other Activities of CBS Related to Agricultural Statistics

a) Smallholder Estate Crops Pilot Surveys

Attempts have been made by CBS to improve the statistics on estate crops grown by smallholders because the DG of Estate Crops does not have enough personnel in the field to carry out the monthly census of eyewitness accounts so the forms are filled out and returned to the D G sporadically.

In 1976 CBS conducted a pilot survey on rubber smallholders in south Sumatra and Jambi and, based upon the pilot, other surveys were conducted in five provinces (South Sumatra, Jambi, Riau, North Sumatra and West Kalimantan) in 1977/78. Pilots were also conducted on sugar cane, coffee, tea, pepper and coconut.

The results from the pilots were considered to be unsatisfactory because the sampling frame, a list of farms from the 1973 census, was out of date and the census list itself was not complete because the areas outside Java had a sample census. 1/

b) Livestock Survey

In 1980 three pages of questions on livestock were added to the annual farm household (cost structure) survey (refer to next section). The questions dealt with an inventory of livestock by type, age and sex, number slaughtered, birth rate and death rate.

An effort will be made to obtain a directory of livestock enterprises dealing with slaughter, milk production and egg production for construction of a sampling frame for surveys and case studies.

c) Farm Household Survey

The Farm Household Survey is also known as the agriculture survey or the cost structure survey because, aside from basic data on land use, tenancy and production practices, data is collected on the cost structure of farm household enterprises. 2/

The sampling frame used for the survey is based upon the comprehensive National Socioeconomic Survey (SUSENAS) list. Ten percent of the SUSENAS households which work in agriculture are selected for the farm household survey. There are an average of 60,000 farm households in the survey and it has been taken annually since 1972. (Note: the SUSENAS survey itself is taken about once every three years.)

The Farm Household Survey provides a vehicle for studying different types of agricultural data through the addition of questions on special topics of interest to the DOA. (The addition of livestock questions in 1980 is an example.) Since it is linked to SUSENAS (either as a module or a separate survey) the data collected can be cross tabulated with socioeconomic data collected in SUSENAS. The only drawback of using SUSENAS as a sampling frame is that any problems related to SUSENAS coverage will carry over into coverage problems for the Farm Household Survey.

1/ For an explanation of sampling frame problems refer to Section III A.

2/ Food and Agricultural Statistics in Indonesia, Unpublished mimeograph prepared for the Eighth Session of the Asia and Far East Commission on Agricultural Statistics in Nepal, Oct. 1980. p.7.

d) **Price Statistics**

The Price Statistics Division of CBS is responsible for collecting data on farmgate prices received by farmers for their production and the prices paid by them for labor, tools, fertilizer and other farm inputs. Those data are collected through interviews of farmers in selected villages. Retail prices for nine categories of nonfarm purchases are also collected in rural markets as well as retail prices in urban areas and wholesale prices at production centers.

e) Food Consumption

Food consumption statistics are gathered by CBS as a component of the National Socioeconomic Survey (SUSENAS). Consumption of nonfood items and services are also collected. SUSENAS has been taken approximately once every three to four years since 1963.

f) Remote Sensing/Area Sampling Frame Construction

CBS is the executing agency for a USDA project for remote sensing and the construction of an area sampling frame based upon aerial photography and topographical maps. A test frame was developed and a pilot taken in Lepan in Southern Sumatra and another frame is being constructed in West Java for a second pilot.

An area sampling frame, once constructed, provides the most up-to-date efficient frame for area and production statistics because the basis for sampling is the land itself rather than a list of farmers' names. The area frame will be stratified according to land use and, although updating (segment rotation) is required approximately once every five years, the frame is more easily maintained than a list frame which is considered (by most USDA experts) to be out-of-date almost as soon as it is compiled.

The constraints to area frame construction are usually the lack of good aerial photography or maps and the costliness in terms of time and money in obtaining them. Furthermore, in the opinion of the Census Bureau, area frames are less efficient than list frames for gathering data on demographic variables, farmer behavior and socio-economic conditions in rural areas. In the case of Indonesia this may not be a problem because the National Socioeconomic Survey already exists and fulfills the need for such data.

B. Department of Agricuture

Given that most of the agricultural statistics collected in Indonesia are processed by CBS, the principal focus of this section of the paper is to discuss the primary data collection and processing done by DOA with little or no collaboration from CBS.

The Department of Agriculture (DOA) is made up of five Directorates General (food crops, estate crops, animal husbandry, fisheries and forestry)

the Agency for Agriculture Education, Training and Extension,1/ the Agency for Agricultural Research and Development (AARD), the Inspectorate General and the Secretariat General (Figure 1). The Agricultural Development Planning and Administration Project is being managed from the Bureau of Planning (BOP) of the Secretariat General and the computer acquired under the project will be installed in the Data Processing and Statistics Center of the AARD for use by all of the Directorates General (DG) of the Department.

The DGs and the AARD are the most important entities in the DOA which collect primary and secondary data concerning agricultural research and production. The data collected and processed by each of them will be discussed in turn.

1. Directorate General of Estate Crops Data Collection

Data Collection

The principal primary data collection activity of the DG of Estate Crops deals with basic information on area, production, inputs, pests diseases and marketing activities for estate crops grown on farms of less than 25 hectares (smallholdings). The data is gathered monthly by mantri kebun from local village chiefs, just as the area statistics on food crops are gathered. The completed forms are sent from 3,200 subdistricts to Jakarta for processing by the Subdirectorat of Surveys and Statistics of the DG.

Tentative plans are being made to gather additional information on farm management and infrastructure but they are a function of the recent reorganization and are still implicit.

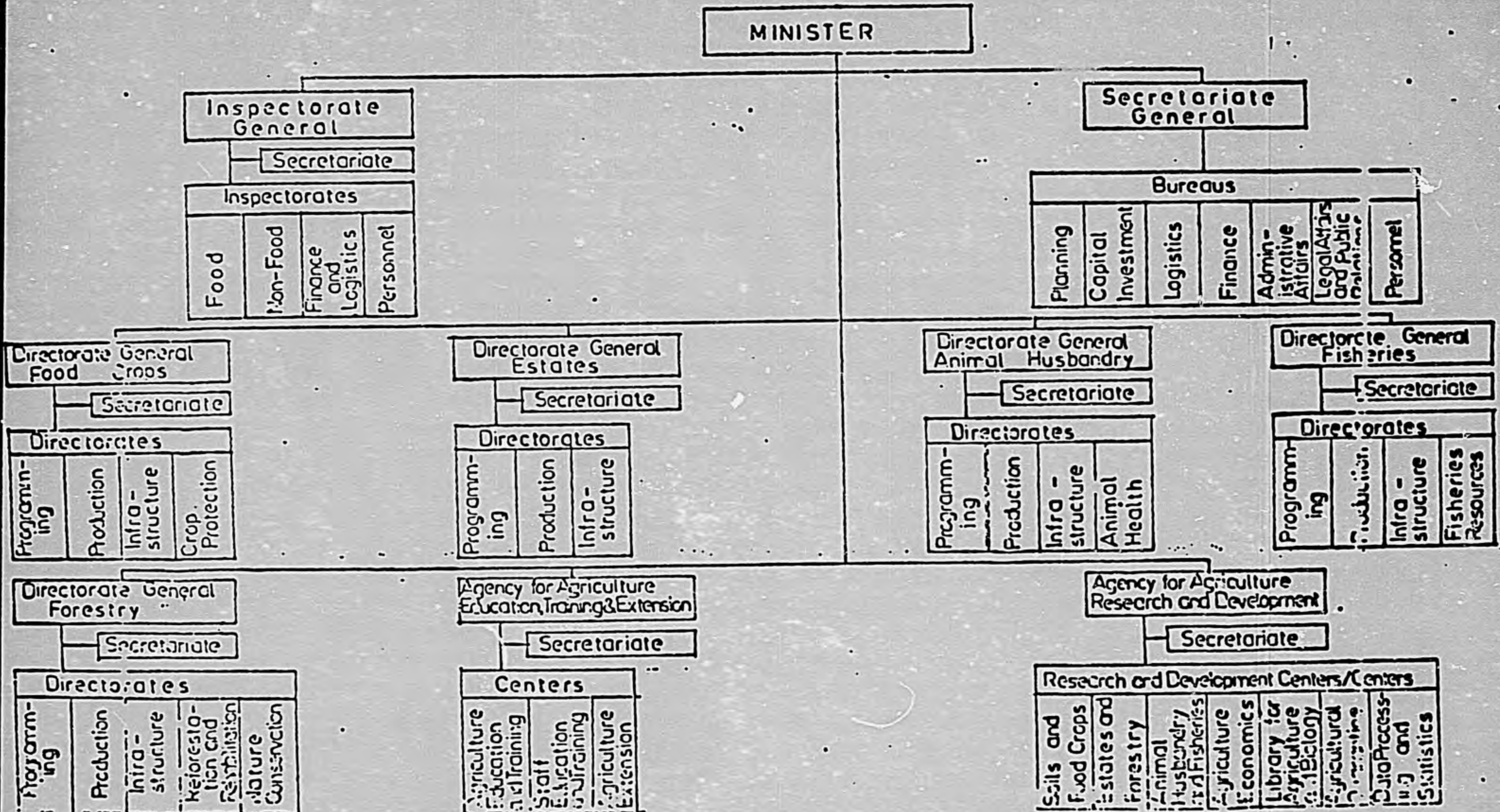
Approximately fourteen people in the Subdirectorate are responsible for compiling reports based upon the primary data gathered on smallholders and secondary data from other sources like the Central Bank, the largeholder statistics from CBS, the Export-Import Bank, the Department of Trade and Labor, etc.

Data Processing

Unfortunately, the system of processing the smallholders data collected by the DG of Estate Crops makes the ultimate report of these data even more tenuous than it might otherwise be. At the present time these data are processed over parallel but partly independent paths within the DG of Estate Crops. One of these paths involves exclusively manual processing, the other involves computer processing. The DG of Estate Crops representative with whom we spoke is responsible for reconciling these two sets of figures thus produced and he expressed dismay over discrepancies as great as 500% (in either direction). It was apparent that he distrusts the computer calculations, probably with just cause, but he continues to work with these results under orders from his superiors.

1/ We could not determine the function of the Agency for Agriculture Education, Training and Extension which appears in Figure 1. It is therefore not discussed in this report.

DEPARTMENT (MINISTRY) of AGRICULT



1) Manual Processing of Smallholder Data

Forms completed at the subdistrict level are manually edited, aggregated and transcribed at the regency level and again at the province level before being sent to the DG of Estate Crops office in Jakarta. A staff of 13 persons works in the central office to perform the final aggregation and produce the required reports. The director of this staff believed that its size should be doubled to cope with the current demands.

2) Computer Processing of Smallholders Data

As part of a larger scheme for computerization undertaken by the DG of Estate Crops in 1977, with the assistance of a U.S. based software firm (COMARC Design Systems), a new form for estate crops data reporting was developed and a computer program was written to aggregate the data collected on these forms in a single report.

In the year that COMARC worked with them, a DG of Estate Crops committee of 10 persons met on several occasions with staff from COMARC to develop forms for recording data in such disparate areas of interest as finances, projects, inventory and estate statistics. The form developed for estate statistics unfortunately was designed to record data on all types of estates, government, private and smallholders. The result was a multipage form that is far too complex for successful completion by DOA personnel at the subdistrict level where it is intended to be filled out. Estate Crops personnel now struggle along using staff at the regency level to fill in the gaps left by the less qualified subdistrict personnel. From the comments on the difficulties being experienced with the new system it seems apparent that the system is not working and is not likely to work as it stands.

We were unable to examine the COMARC prepared COBOL programs for report preparation. They are currently in the hands of a Jakarta contractor (DKI) that we are told will not be eager to relinquish this job. The programs were, however, described to us by Dan Peterson (ISU[1] data processing contractor) as extremely straightforward in design. We assume that the processing consists of little more than basic aggregation of the recorded data. At the present time, forms completed at the regency level are sent to Jakarta where they are handed over by the DG of Estate Crops to DKI for keying and processing.

2. Directorate General of Fisheries

Data Collection

All data dealing with marine and inland fisheries are collected by the DG of Fisheries with no assistance from the mantri statistik of CBS. Production statistics are collected monthly, by mail from marine

[1] ISU refers to Iowa State Univeristy.

enterprises which produce for export; weekly, by fishery agents from marine landing centers; and quarterly, through household sample surveys of marine fishermen who do not produce for export. Inland fishery production data are also gathered quarterly through household sample surveys.

More detailed socioeconomic sample household surveys of fishermen are carried out annually for specific regions and types of fishing on a rotating basis:

1975	marine fishing	- North Coast of Java,
1976	marine fishing	- Malacca Straits,
1976	marine fishing	- Macasa,
1977	brackish water fishing	- Java,
1977	brackish water fishing	- South Sulawesi.

Plans are being made to proceed with fresh water fishing in Java.

Eighteen people work in the subdirectorate of Statistics of the Planning Directorate of the DG. They are in charge of the data collection activities which were designed during a five-year period under a UNDP/FAO project.

According to the people interviewed, the biggest problem encountered in collecting fishery statistics is the lack of sufficient funds to pay the fishery agents who are responsible for taking the household production surveys.

Data Processing

The DG of Fisheries in conjunction with the CBS uses computers in the processing of its collected data. The lowest level of resolution available for DG of Fisheries data is the regency level where personnel in the DG of Fisheries offices at this level aggregate questionnaire responses and record data for the surveys conducted by this DG. The recorded data subsequently passes through offices at the province level to the Jakarta headquarters where it is reviewed before being handed over to CBS personnel for keying and processing. CBS processes the following surveys for the DG of Fisheries.

(1) A very straightforward survey on prawn production involving approximately 2000 documents is processed annually. Questionnaires completed by prawn fishermen are gathered by the DG of Fisheries and handed over to CBS for processing. CBS considers this a small job requiring no more than 7 CPU hours on the ICL 2904 system at CBS.

(2) A second DG of Fisheries survey involving approximately 4500 documents, each of 10 pages in length, is processed annually. This project is a considerably larger effort. The forms used were not designed for computer keying and there are COBOL validation programs for checking the keyed data. Approximately 100 basic tables are required. The coding of these tables is in COCENTS.

(3) A more comprehensive survey designed with the help of Dr. Yamamoto (an FAO advisor) between 1973 and 1978 is intended to be processed annually although it has not been done for some time. Seven different forms containing a recap of data from each regency are prepared by DG of Fisheries personnel in the regency offices. These forms are gathered by the DG of Fisheries office in Jakarta and they are responsible for setting a cutoff date for return of the forms and turning the data over to CBS for processing.

A production run of this survey requires approximately 6 person months of programmer time and about 25 CPU hours on the ICL 2904 processor. There are no members of the DG of Fisheries staff in Jakarta with experience in data processing or knowledge of the programs being used for processing their survey data though an interest was expressed in training both in survey methods and data processing.

3. Directorate General of Forestry

The major record keeping activity carried out by the DG of Forestry is the inventory of trees on the government owned forest complexes. This has been done since 1930 and requires no modification. The DG of Forestry has maintained a computerized system for forest inventory since 1971. The CBS assisted in setting up the computerized procedures they now follow.

Each item in the forest inventory is recorded in the field by DG of Forestry personnel on a form designed for keying to a 22-character computer record. At this time the data is being keyed to diskette using a Q1 micro computer in the DG of Forestries offices in Bogor. The current rate of keying stands at 14,000 records per month though there have been times when receipt of these records has reached a rate as high as 100,000 per month. There are approximately 400 (single density) diskettes in the forest inventory library.

There are a number of PL/1 programs designed for use on the Q1 system to work with this data. Part of the 22-character record is a local tree name which must be converted into one of more than 4000 trade names and a program has been written that uses a dictionary of names to perform this conversion. There are programs also used to generate reports.

Two DG of Forestry staff members have had experience in computer programming and there is an aggressive interest within this DG in developing computerized data systems.

4. Directorate General of Animal Husbandry

The data collection system implemented by the DG of Animal Husbandry is similar to that of the DG of Food Crops and the DG of Estate Crops. Data on animal inventories is collected monthly at the subdistrict level by mantri hewan and passed through the district and province levels for review by field personnel of the DG. It is then sent to Jakarta for final processing by hand, publication and report writing. Other information (12 forms in all) is gathered routinely on a monthly and quarterly basis and published annually.

The most important bottleneck in the data collection and processing is at the subdistrict level because there are fewer mantri hewan than there are subdistricts so the workload is very heavy and forms are often late or never filled out and sent to Jakarta.

Occasionally special surveys are carried out in conjunction with universities and CBS.

The DG of Animal Husbandry relies on manual processing of 12 simple forms to obtain the information it requires. As in other Directorates General, data collected and recorded by the DG's mantri at the subdistrict level are reviewed, adjusted and aggregated by their personnel at the regency and province levels. A staff of 5 persons in the Jakarta offices perform the final review and aggregation of the data.

The DG of Animal Husbandry has worked in cooperation with the CBS on surveys that have involved computer processing but the staff has not been exposed to computers and has no computer processing experience.

5. Bimas (Bimbingan Massal) 1/

The BIMAS program was initiated in the early 1960's and is still in existence with a broadened scope of work which includes the promotion of the use of high-yielding varieties of rice, riceland intensification programs and evaluation of rice programs.

The principal data collection activity is done in tandem with the data gathering of the DG of Food Crops and CBS. Mantri Pertanian cooperate with BIMAS field agents to collect the special information required by BIMAS on a monthly basis to monitor its rice intensification program. This information is collected in addition to the more basic information in the food crop data collection program of the DG of Food and CBS. The special information required by BIMAS is sent to Jakarta every month for manual processing by 5 people in the central office of BIMAS.

Bimbingan massal means literally "mass guidance" and refers to a government program to promote the use of high-yielding varieties of rice by using senior university students who live with and teach farmers in the villages.

Annual statistics are collected on the potential for intensification. The design of this program was done in collaboration with the Institute Pertanian Bogor (IPB, the Agricultural Institute in Bogor). The form was designed for computer processing and contains 33 80-character records. Processing is done by contract with a private firm recommended by CBS. The facility currently being used has IBM equipment.

Occasionally (about twice every five years) surveys are taken to evaluate the BIMAS program. The questionnaire used was designed in conjunction with CBS for the first time in 1979 and improved for easier processing in 1980. The results from the 1980 survey have still not been processed; there are 3200 questionnaires of 48 pages each. CBS processed the 1979 survey for BIMAS, taking five months to complete the job. At that point a BIMAS staff of 15 persons used a total of 500 person-days to complete the report required of them. The data collected in 1980 have not been processed yet because CBS is presently involved in work from the 1980 population census.

The BIMAS representative with whom we spoke expressed an eagerness to develop data processing capability within his staff and was willing to give two members of his staff up for training and also, if necessary, to add as many as two people to his staff for keying.

6. Agency for Agricultural Research and Development

The particular mandate of this Agency precludes its involvement in the collection or processing of routine agricultural statistics. Furthermore, from our limited exposure to the work of AARD we can only surmise that the numerous ongoing projects under the Agency's auspices are under the supervision of individual researchers.

The only project of which we were able to gather specific information was an extensive forestry survey on wood consumption being conducted in 11 provinces outside Java. The survey being conducted by CBS personnel involves a total of 368 census blocks or 36,800 households. The 70 page questionnaire is extensive and includes 9 large tables. CBS prepared a manual for data collectors and will be responsible for keying the data. Present plans are for the data to be processed in Australia under the supervision of the survey director.

We would expect the computer processing demands of AARD to increase steadily with the opening of its computer facility later this year, expecially if SPSS should become available.

III. Assessment of Institutional Capability to Modify the Present System of Agricultural Statistics

A. Improvement of the Quality and/or Scope of Agricultural Statistics

1. Major Complaints

Dissatisfaction with the present system falls into the following categories: scope, accuracy, timeliness and duplication.

The scope of the present system of Agricultural Statistics is criticised because there is increasing interest in DOA in a new policy of crop diversification. In the past, the overwhelming concern was with the promotion of the use of high-yielding varieties of rice and intensification of land use in rice through increased use of inputs ranging from fertilizer to irrigation. Therefore the present system of agricultural statistics has a strong concentration of statistics for rice production and is weak in potentially high value crops like sweet potatoes, mung beans, vegetables, pepper, vanilla, etc. which, if produced, might reap a higher value of production from the land and increase income in rural areas. Other statistics which would broaden or improve the scope of agricultural statistics are modification of the current land classifications used (irrigated vs. not irrigated and technical vs. nontechnical), improved data on wages, mobility across agricultural subsectors, prices of production, inputs and nonfarm commodities.

The accuracy of the statistics produced is questioned because of the range of variation within statistics for smallholder estate crops, the difference in results between estate crops statistics edited and tabulated by hand and by computer (which is said to be as much as 500% for some items) and the difference in statistics produced by DOA and CBS which go through two different channels of processing.

Many of the agricultural statistics series published by CBS are one to three years behind schedule. This is undoubtedly the reason why DOA passes the data from the field through its own channels of processing at the district and province levels and has access to the data for internal use before it is officially released by CBS.

There is presently duplication of effort between CBS, DOA and BIMAS in the production of agricultural statistics. This is because by law CBS is responsible for the collection of basic statistical data but the DOA and BIMAS often require data which in some cases is slightly different and in other cases totally different from what is contained in the official series published by CBS. The scope and degree of duplication is indeterminate due to the lack of a detailed inventory of the types of data collected.

2. Methodological Problems and Recommendations

a. Scope of Present Agricultural Statistical System

A quick review of the types of data collection, the frequency of collection and the geographic coverage (shown in Table 1 of this report) in conjunction with reports written by personnel from DOA and CBS (which are referenced in this report and attached as appendices) indicates that the scope of the system is comprehensive when compared to what is generally considered to be an adequate system of agricultural statistics. 1/ Benchmark data on physical resources such as land utilization, ground and surface water supplies, forest and range resources and transportation are collected by BIMAS, CBS and the DG of Forestries. Benchmark data on human and economic resources are collected by CBS in the population census and surveys and the National Socioeconomic Survey (SUSENAS). Farmer/producer behavior information is collected by BIMAS in its rice project evaluation surveys and CBS in its Farm Household Surveys. Agronomic data showing how major crops respond to management inputs are collected in case studies done by AARD and IPB (Institute Pertanian Bogor). Current supply data including production, area and yields for most important crops are collected by CBS, the DG of Food Crops and the DG of Estate Crops in their regular series. Prices are collected by the Department of Trade, CBS and the Bulog (Bureau of Logistics).

What is desperately needed at this point is a comprehensive inventory of the data being collected because although a quick review indicates that the system comprises the basic information required, it is extremely difficult (1) to understand how CBS, DOA and BIMAS interact to produce the body of data available, (2) to determine exactly what variables are collected for what types of crops and livestock at what frequency, (3) to determine how much duplication of effort is taking place and (4) to determine what data is available for users and in what form (tables, tapes, etc.).

Such an inventory should be useful, if not necessary, for cutting back on duplicative efforts and creating a data base management system.

1/

a) Minimum Information Systems for Agricultural Development in Low-Income Countries: by Ralph W. Cummings, Jr., Seminar Report No. 14, The Agricultural Development Council, Inc. New York, Sept. 1977.

b) Statistics and Studies for Agricultural Development Planning: A Phased Program, Agricultural Planning Studies No. 10, FAO, Rome, 1968.

We, therefore, recommend that CBS be contracted to develop an inventory of agricultural statistics because, on behalf of the Government, it acts as a coordinating body of all statistical activities of the other government agencies. We also believe that it is within reason to suggest that the effort be paid for out of the loan funds from the Agricultural Planning and Administration project because one of the major purposes of the project is to improve the data management system and any such effort will require an inventory as a foundation for its development.

The effort will require two to three person months of time by an individual who speaks Indonesian and is already familiar with the present system. It will also require cooperation from the DGs of DOA which are responsible for the data collection in many areas of agricultural statistics. This, in turn, will require agreement by the committee for agricultural statistics in BAPPENAS and the Minister of Agriculture.

The inventory would include the items noted in Table 1 but be expanded to include the types of variables collected for each type of crop and livestock, the proportion of forms actually used in the tabulations (rate of response by form used) and the agency responsible for its collection, processing and publication. It would also be helpful if the inventory were translated into English for use by foreign advisors and institutions.

We also recommend that, rather than looking for ways to increase the scope of agricultural statistics, efforts be made to set priorities for the improvement of the accuracy and timeliness of the data which are already being collected. Any requests to increase the scope of statistics should be evaluated in terms of the system as a whole (within the agricultural statistics committee) so that trade-offs among scope, quality and timeliness are considered and the necessary funds and personnel for changing the system are made available.

b. Frequency of Collection

Most of the data collected on crop area and yields, livestock inventories, forestry inventories and fisheries production are collected very frequently - at least monthly if not weekly. Given the nature of the agricultural production in Indonesia (continuous planting and harvesting for many crops, continuous fishing and large forested areas) we cannot recommend that the frequency of collection be cut back for most items. 1/ Some consideration might be given, however, to reducing the frequency of collection of data on livestock and some perennials which have a low rate of household consumption.

c. Statistical Methodology (Sampling vs. Census)

One of the reasons why accuracy and timeliness are not optimal is that the principal data collection activities are censuses rather

---

1/ Reduction of the frequency of collection should not be precluded. The issue requires closer study.

than surveys and the amount of work required is overwhelming and hard to control because so many field agents are involved. Ideally, all of the continuous statistics on area and production of crops should be collected through sample surveys. The major constraint, however, is the lack of a sampling frame. Note that sampling was attempted with unsatisfactory results in the pilot surveys of smallholders in estate crops due to the inadequacy of the frame. The sample surveys which are done rely upon a list of households based upon the SUSENAS survey which may have its own sampling frame problems (Table 1).

The lack of an adequate sampling frame for agricultural area and production statistics is one of the most intractable problems confronted by most developing countries interested in improving their agricultural statistics. It is intractable not because methodologies do not exist to construct frames but because the resources in terms of time and money are usually not available for such investments.

Two kinds of frames, if available, could be used: a list frame or an area frame. A list frame is a list of the names and location of all farmers in the country which could be obtained through a population census or an agricultural census. It is possible to develop a list frame without having to rely on an up-to-date listing of farm households: first stage areas could be chosen, like villages or blocks within villages, using some convenient measure of size, like counts from an old census; then lists need only be constructed for second-stage segments chosen within the sample first-stage units. This is probably the method used in SUSENAS and the other farm and fishing household surveys.

An area frame is constructed by delineating, as precisely as possible, agricultural areas (or any area) based upon aerial photography or good topographical maps with consistent scales. The frame is stratified according to land use and segments are chosen for enumeration.

The problem with list frames is that they must be kept up-to-date by relisting (either in all areas, or the second stage segments chosen). The problem with area frames is often the lack of good photography or maps and the time required to obtain them and then delineate the frame. Area frames must also be updated by restratifying for land use and rotating segments. However, once constructed area frames are considered to be more efficient for continuous area and production statistics.

The cost in terms of time and money involved with constructing and maintaining either frame in a country the size of Indonesia with so many inhabited islands may be prohibitive.

A pilot project for the construction of area frames is already underway at CBS. If the results of the pilot surveys based upon the frames are good, consideration could be given to developing the frame nationwide or at least for the most important agricultural areas. Before such consideration is given, an estimate of the time and cost required should be made for both construction of the frame and implementation of the survey.

The population census, if it listed variables which determine which households are farm households, could provide a list frame. The upcoming agricultural census if done with full coverage on all islands could also provide a frame. (Note: by full coverage we mean visiting all households to create a basic list -- not asking all census questions of all farm households.) The samples selected within one to two years after the census will be more efficient than those selected later and relisting should be done once every three years (no less than five years).

A survey program based upon either frame should not replace the present system until pilot results are carefully examined and compared to results of the present system. It is possible that the present system is the best given the resources available.

d. Quality Control in the Field

If it proves to be too costly to employ sampling it will be difficult to significantly change the manner in which data are presently collected on area and production of crops. Such changes would probably involve conversion to a system based upon interviews, more objective measures of area or a combination of both.

Working with the eyewitness systems that presently exist for area and animal inventories, the changes which can be recommended are marginal. For example, of the total number of forms which should be returned to CBS for area statistics from the subdistrict level only 60% are returned and 80% of these are from Java. The reasons suggested for this low return rate are that there are not enough mantri pertanian at the subdistrict level in some parts of the country, they are paid from provincial level funds which may not be forthcoming for statistical work, they are required by DOA to fill out other administrative and statistical forms which may be given higher priority by DOA and they are not dedicated to the collection of statistics alone.

Since the return rate is low and the biggest bottleneck to data processing is late receipt of forms in the central office, the chief of the Agricultural Statistics Division of CBS is proposing that a survey be made of the problems of the mantri statistik and the mantri tani at the subdistrict level to find out what the problems are. We support that recommendation and suggest that the survey be directed at finding out the number of mantri available by subdistrict, their workload, the return rate of forms, the problems they encounter in filling out the forms and the suggestions they have for improving the timeliness of the reporting procedures.

From what we have been able to learn, the only ways to improve the quality of field work for the eyewitness system are to increase the number of mantri in areas where they are scarce, raise their pay and status, remove from their scope of work matters not related to statistics (in other words, make them all mantri statistik), establish a national priority as to which forms should be returned first, pay them with central funds, cut out duplications of effort, increase the quality and amount of training and supply them with pencils, paper and hand calculators.

e. Data Processing

The approach to processing the primary data currently used in the CBS is labor intensive and apparently entirely adequate to meet the current needs of the CBS. The fifty-member staff currently responsible for processing collected data on crop area, crop yield and crop production is apparently well trained and experienced in its work and is capable of performing the editing and aggregations required in a timely manner. Because the aggregation procedures are straightforward, the manual processing can be done accurately and quickly.

The quantity of data being collected and processed in this manner annually is sizable. We estimate that it would require a minimum of 20 person-years to accomplish the keying of the data being received annually. The consequence of a move to computerized processing would be the ultimate replacement of fifty persons exercising a fair degree of responsibility in their work by twenty persons exercising a far lower level of skill.

In addition, several person-years of a highly skilled (and also scarce) native programmer would be required to provide for the development, maintenance, and operation of these computerized procedures.

Such a conversion should not be undertaken before serious thought is given to the overall data collection and processing effort or before a manifest need for more complex processing methods urges the introduction of computer processing.

3. Practical Constraints

In order to make a significant improvement in the quality and timeliness of agricultural statistics it will be necessary to establish priorities and make available the funds and personnel required to implement sought after changes.

a. Institutional Requirements

The Committee on Agricultural Statistics of BAPPENAS should be responsible for hearing and analysing criticisms and suggestions related to the scope, quality and timeliness of agricultural statistics.

FORM	FREQUENCY	NO. FORMS RECEIVED ANNUALLY[1]	APPROXIMATE NO. CELLS PER FORM	NO. BYTE OF DAT GENERATI ANNUALLY (IN MILLIONS OF BYTES)	MAN HOURS[3] REQUIRED FOR KEYING ANNUALLY	MAN YEARS[3] REQUIRED FOR KEYING ANNUALLY
SP IA LUAS TANAMAN PADE	MONTHLY	39,000	120	18.7	3120	1.8
SP IB LUAS TANAMAN PALAWIJA	MONTHLY	39,000	240	37.4	6240	3.6
SP II LAPORAN TANAMAN SAYVE-SAYURAN	MONTHLY	39,000	150	23.4	3900	2.2
SP III LAPORAN TANAMAN BUAH-BUAHAN	QUARTERLY	13,000	180	9.4	1560	0.9
SP IV LUAS SERANGAN HAMA PENYARIT DAN BENCANA ALAM	MONTHLY	39,000	340	53.0	8840	5.1
SP VA LAPORAN PENGGUNAAN TANAH	ANNUALLY	3,250	75	1.0	160	0.1
SPVB LAPORAN ALAT-ALAT PERTANIAN	ANNUALLY	3,250	100	1.3	216	0.1
ESTATE CROPS I	MONTHLY	14,400	150[4]	8.6	1440	0.8

FORM	FREQUENCY	R[illegible] A[illegible]Y[1]	APPROXIMATE NO. CELLS PER FORM	GEN[illegible] ANNU[illegible] (IN MILLION[illegible]S)	REQUIRED FOR KEYING ANNUALLY	REQUIRED FOR KEYING ANNUALLY
ESTATE CROPS II	ANNUALLY	1,200	1,200[4]	5.8	960	0.5
DAFTAR II KETTRANGAN HASIL UGINAN	CONTINUOUSLY	80,000	150	48.0	8000	4.6
TOTAL				206.6	34436	19.7

1
Based on full response from 3,250 kecamatan officer.

2
Based on estimated 4 characters/item, 1 character/byte.

3
Based on estimated 4 keystrokes/item, 12,000 keystrokes/hour, 1750 work hours/year, but including 100% verification.

4
Estimate of 150 cells/page of questionnaire.

It should establish priorities for their improvement; for example, it should choose between improving area statistics or increasing the data generated on high-value crops. (From what we could gather it is already performing this function) A national mandate is absolutely necessary for making an impact upon the system of agricultural statistics because personnel and funds must be made available for carrying out recommendations.

The Committee should also be responsible for coordinating the relations between CBS and DOA so that redundancy of work is not permitted. For example, both CBS and DOA are considering the development of a computerized agricultural data base.

Any program for the improvement of agricultural statistics should be implemented through CBS with the cooperation of DOA because CBS has a considerable amount of experience in sample survey and census design and implementation and automatic data processing. Personnel in CBS are aware of the problems with the present system and on many fronts are anticipating possible means of improvement through pilots of sample surveys, area frame pilots, a survey of the problems of the mantri, coverage of the upcoming agricultural census and its use or the use of the population census as a sampling frame.

Although the DOA is greatly responsible for collection of agricultural data in the field it has little or no experience in sampling, survey and census design and implementation and automatic data processing. It should take a secondary role in the area of agricultural statistics for the time being.

b. Resource Constraints

Any project or program aimed at improving agricultural statistics would be a multi-million dollar program and would require five to ten years to implement. Marginal improvements in the quality of field work under the present eyewitness system would also be costly because more mantri would have to be hired and given higher incentive pay and a more narrowly defined scope of work.

B. Improve Ability of DOA to Collect and Process Primary Data

1. Rationale

If there is a sincere desire within the DOA to improve its ability to collect and process primary agricultural statistics under the auspices of the Agricultural Planning and Administrative Project, training should be given in sampling and survey statistics as well as data processing. With formal and on-the-job training it might be possible for DOA to (1) improve some of the statistical programs under its control in the areas of forestry, fisheries and livestock (2) transfer some of the functions presently carried out by CBS to DOA and/or (3) design and implement special one-time surveys which will fulfill special ad hoc requirements for project and program planning within the DOA.

2. Types of Training

a. Sampling

Training in sampling should cover basic statistical theory, simple random sampling, simple stratified sampling, simple one and two-stage cluster sampling, stratified single and multi-stage sampling, ratio estimation, difference and regression estimation, double sampling and sampling for time series. The student should also be taught variance estimation techniques designed to simplify calculation and reduce the cost of variance estimation for more complex sample designs such as ultimate cluster estimates, random group method, Mc Carthy's half-sample replications, Tukey jack-knife replications and Taylor's series approximations. Emphasis should be given to applications of these techniques to practical problems relevant to a developing country using simulated problems and data. Instruction might also be given in frame construction for agricultural surveys and objective measurement of area and yield.

b. Survey Statistics

Training in survey statistics should introduce the student to the major steps involved in the initial process of designing a census or survey including specification of objectives, consideration of budget and sample limitations, decisions about which variables should be measured and how and selection of methods for collecting and processing data.

Training should continue in the details of survey and census implementation; use of maps, questionnaire design, quality control in field operations, principles of editing, coding and imputation, specification of editing instructions and tabular output, verification of tabular results, etc.

c. Statistical Data Processing

Training should include introductory courses in computer processing fundamentals and fundamentals in programming as well as programming in COBOL and/or FORTRAN. Aside from learning how to program, the student should be introduced to design of surveys and censuses and statistical methods. More detailed instruction should be given in design of tables and questionnaires, editing, coding and imputation procedures. Courses in tabulation and editing packages as well as data base management are optional. The purpose of the training is specifically directed at training data processors to design and program software systems to process primary data. (For more information about the kinds of processing being referred to consult section 5C of this report, "On-the-Job Training in Data Processing").

d. Short-Term, Long-Term and On-The-Job Training

Short-term training involves short courses of two weeks to six months which are useful as refresher courses for those with basic knowledge of the subject matter or for beginners who will be guided closely by senior staff in on-the-job training after they take the courses.

Long-term training involves courses or programs of one year or more and may lead to a masters degree or Ph.D. Such training is required to build a core of experts who have been chosen as potential senior level personnel to form the basis for a statistical program.

On-the-job training involves teaching a skill through actual work implementation. It is usually necessary for insuring that those who have been formally trained can exercise the functions for which they were trained and requires that a manageable task be specified for their training.

3. Requirements for Successful Training and Absorption of Trained Personnel

a. Central Control and Design of a Purposive Training Scheme

In order to design an intelligent training program which will meet the needs of DOA in the areas of primary data collection and processing it will be necessary for the Agricultural Information Advisory Committee of DOA to determine:

(1) the number of people to be trained in each area,
(2) the distribution of people to be trained among the DGs and AARD,
(3) the amount of time each trainee can be allowed to train without disrupting the ongoing work of the DG or AARD,
(4) the amount of financial resources they want to commit to each type of training, and
(5) which potential primary data collection activities (being carried out or planned for implementation) could be used as vehicles for on-the-job training.

b. Calendar Time and Organization of Training

It will require at least five years for the completion of basic training in data collection and processing. This assumes that those people who have acquired one to two years of long-term training will proceed with on-the-job training upon their return to work and will be assisted by those who have received short-term training. On-the-job training should ideally include technical assistance by CBS advisors or foreign experts so that problems encountered in the initial work undertaken can be smoothly resolved.

Since on-the-job training is critical, it is necessary to specify activities in DOA which require and would benefit from expertise in the areas of sampling, survey statistics and statistical data processing. If no such activities can be agreed upon by the Steering

Committee as deserving attention, then there is no need for training in the first place.

4. Suggested Training Program

We suggest that seven people, one from each of the five DGs, one from BIMAS and one from AARD, be trained in agriculture survey statistics for a period of one year at the International Statistical Programs Center (ISPC) of the Census Bureau in Washington, D.C., that two people from AARD be trained in sampling at ISPC or any university chosen by the Committee and that two programmers from AARD be trained at ISPC for a period of one year in programming and statistical methods.[1] (Refer to Table 3). The six people trained in agriculture survey statistics[2] would serve as resource persons in the area of primary data collection for their respective DGs and as liaisons between the DGs and the samplers and data processors in the Data Processing and Statistics Division of AARD. The samplers would assist the DGs in the design of samples for their survey activities and the programmers would assist the DGs in development of software systems by programming themselves or assisting programmers provided by the DGs. This entire group of ten people would provide the core of experts which could form the basis for improving primary data collection and analysis in DOA.

In conjunction with experts at CBS or foreign experts, this core of trained people could develop a series of two to four-week courses which can be taught to personnel in the planning bureaus of the DGs in relevant areas like questionnaire and forms design, development of training manuals for field work (interviewers and supervisors manuals), writing specifications for computer editing and tabulations, editing and computation procedures, systems design for processing of primary data, quality control programs for survey and census implementation, development of training sessions for mantri, etc. The core of trained people should also become involved in planning and implementing methodologies which will improve the data collection and processing activities in the DGs. This on-the-job training might require assistance from experts with more experience to resolve unforeseen problems. It should also involve as many people as possible in the planning bureaus of the DGs so that proper procedures are at least understood by the majority of the people responsible for primary data collection and processing.

Possible data collection activities which might serve as training vehicles are the DG of Fisheries Socioeconomic Survey of Fishery Households, the DG of Estate Crops Census of Smallholders, the DG of Livestock census of livestock inventories, BIMAS survey for the evaluation of rice intensification programs, or any one-time ad hoc survey which is required by DOA but cannot be implemented by CBS when required.

1/ Refer to the training brochure from International Statistical Training Programs, U.S. Dept. of Commerce, Bureau of the Census, August 1980-1981, p.13 for agriculture statistics, p.19 for computer data systems and p.11 for sampling.

TABLE 3: FINANCIAL REQUIREMENTS FOR SUGGESTED TRAINING PROGRAMS FOR COLLECTION AND PROCESSING OF PRIMARY DATA IN DOA (IN U.S. $)

I. LONG-TERM TRAINING[1]		
7 people trained for one year in survey statistics	7 x $11,575	$81,025
2 people trained for one year in data processing and statistics	2 x $11,575	$23,150
2 people trained in sampling for one to two years	2 x $11,575	$23,150
TOTAL LONG-TERM TRAINING[1]		$127,325
II. DEVELOPMENT OF SHORT-TERM COURSES[1]		
12 to 18 person-months foreign technical assistance	18 x $7,000	$126,000
III. ON-THE-JOB TRAINING IN SELECTED SURVEY		
6 to 12 person-months foreign technical assistance (per survey)	12 x $7,000	$84,000
TOTAL TRAINING PROGRAM		$337,325

[1] Excludes international travel and subsistence. A minimum of 70/70 score is required on ALIGU.

5. Description of On-the-job Training

a. Sampling

The purpose of on-the-job training in sampling is to insure that the sampling expert can apply his/her knowledge of sampling principles and procedures and variance calculation to a real-life situation which responds to the needs of the user.

The sampling expert will be required to do an inventory of the mapping materials available for the regions in which the survey is to be conducted and to judge if the quality of maps is adequate for use in selection of enumeration districts (or first-stage areas). He will have to design the sample using stratification that meets the requirements of the user taking care that enough observations are in each of the strata to allow for reliable results. He/she will have to compute weights and design the variance calculations for expansion of the sample and estimation of coefficients of variation. He/she will also be responsible for working with the programmers who will be responsible for programming the variances.

Ultimately the sampling expert will be responsible for documenting the sampling methodology and variance formulas so that future users do not have to guess about the nature of the sample when they analyze the data.

b. Survey Statistics

The survey statistician is the key person in the design and implementation of a survey or census. He/she is responsible for developing the calendar of activities and the resource requirements in terms of personnel and money required to implement the survey. He/she is responsible for coordinating the work of both the sampling expert and the data processor making sure that their inputs are adequate and timely. He/she is, in effect, the manager of the activity.

The survey statistician is responsible for directing the implementation of the data collection activity. This includes:

(1) the design of the questionnaire and all administrative forms required for the flow of questionnaires to and from the field, and through all of the stages of processing

(2) the development of a training program for the interviewers and field supervisors

(3) the development of field manuals for interviewers and supervisors and, if required, training guides

(4) the writing of specifications for the computer and/or hand editing of the data as well as the tabulations to be programmed

(5) the writing of the editor's manual and training of editors

(6) the review of finished tables and assisting the programmers in debugging the programs

(7) the documentation of all survey procedures employed in the implementation of the survey,

The survey statistician will require basic knowledge of all of these steps and must direct other people in carrying out the work to implement them.

c. Data Processing

Data processing personnel should be actively involved with the survey statistician from the very earliest stages of any survey project, to insure that any plans that are made for data collection or data presentation may be made in such a way to facilitate processing. In particular, a data processing representative should take part in the design of the survey questionnaire to insure that all information recorded on the questionnaire can ultimately be used by the keying staff in their work of transcribing the information to a machine readable format. When questionnaires are complex or include many items that may be blank, it may be advisable to use a method which we refer to as source coding. This method involves placing a number on the questionnaire at the position of each possible response. Each item on the questionnaire may then be referred to by a unique number. This procedure often reduces the number of keystrokes required to enter the data on a single questionnaire since those items that are not answered do not need to be accounted for. This also simplifies the work of the keypunch operator by removing the concern for exact column positioning for every response.

The data processing system for producing publishable tables from the keyed survey data will generally involve all of the following:

(1) A computerized check-in procedure. The keyed data records are checked against a master list of all questionnaires. This procedure makes possible the elimination of duplicate entries and the later entry of missing questionnaires.

(2) A computerized routine for file generation. The keyed records are sorted, examined for a number of basic errors, and finally reformatted before being entered into the file that will later be used for table production. Ordinarily the format of the newly generated file will be one that requires a minimum of further machine processing for aggregation (i.e., half or full word binary values are used).

(3) A computerized listing routine. This program enables the user to print selected questionnaires or selected items within questionnaires and it is generally parameter driven.

(4) A computerized update routine. The update program makes it possible for questionnaires to be deleted or inserted in the questionnaire file and for existing items to be changed.

(5) A computerized valid value editing routine. Every value in the newly created file is checked to determine its validity or reasonableness. Each item on each questionnaire must fall within a specified set of values for the given item.

(6) A computerized internal consistency edit. Every record is examined according to a carefully constructed set of rules to reveal inconsistencies in the recorded data.

(7) A computerized routine to recode specified responses and/or to impute values under exactly specified conditions.

(8) A series of computer programs to generate tables summarizing the questionnaire responses. Ordinarily these programs would be written in a special table generation package language (such as CENTS or COCENTS) and would follow specifications prepared by the survey statistician in conjunction with a subject matter specialist.

(9) A computerized routine to calculate variances for selected variables written in accordance with specifications prepared by a mathematical statistician.

A flow chart of a typical system is attached. (Refer to figure 2)

C. Improve Ability of MOA to Process Primary Data

1. Establishment of a Computer Facility within AARD

Under the provisions of a contract between the DOA and Iowa State University a computer center will be established for general DOA use within the Center for Statistics and Data Processing division of the Agency for Agricultural Research and Development. The facility is expected to be operational early in 1981.

a. Hardware

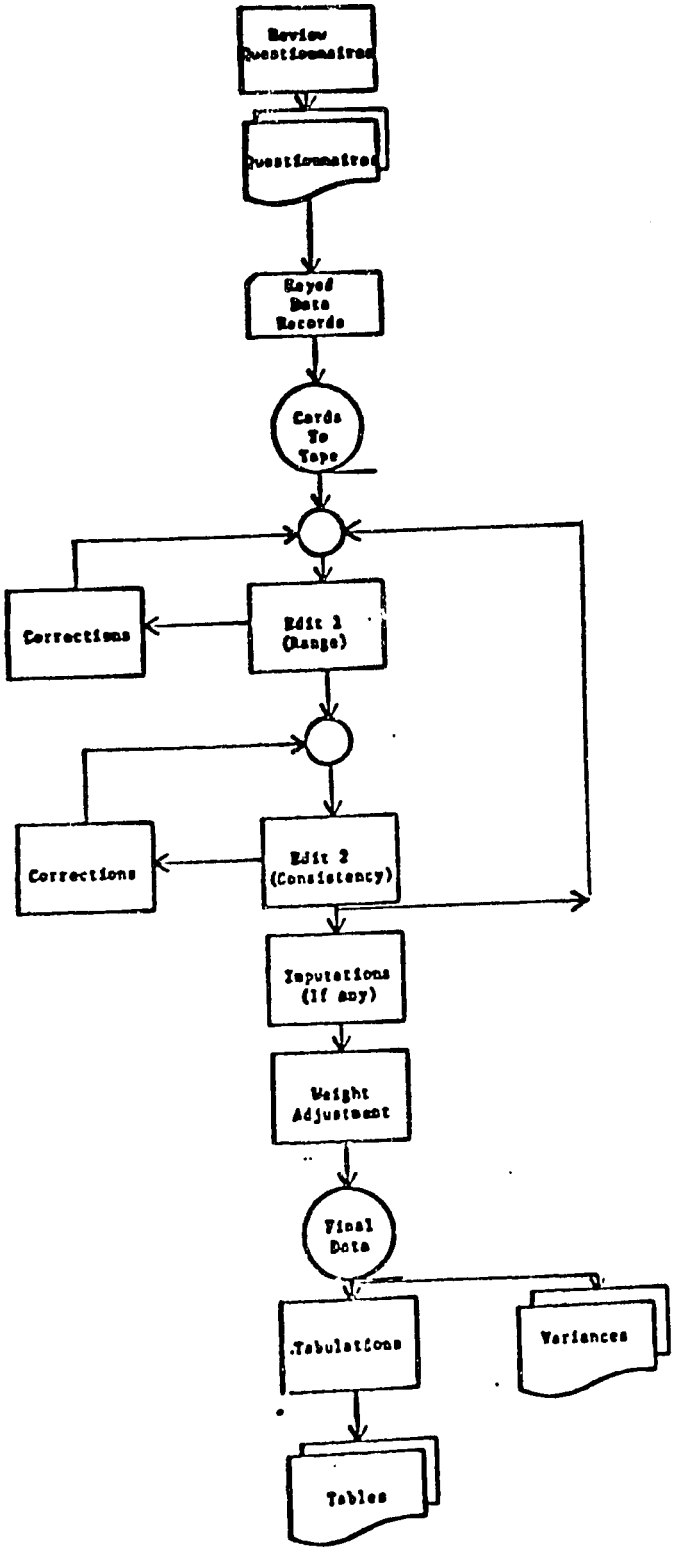
The computer system to be installed includes the following equipment:

a Honeywell Level 6, Model 53 mini-computer including 512 KB memory

4 9-track 800/1600 bpi tape drives (75 ips)

3 256 MB disk storage devices

FIGURE 2. FLOWCHART OF DATA PROCESSING SYSTEM



1 300 lpm printer

4 station printers (120 cpm) and

4 CRT (Cathode Ray Tabs) display units that together form 4 remote stations

1 500 cpm card reader

Under separate contract between the DOA and the same hardware supplier the DOA expects to purchase a Honeywell Mini 6 system and six CRT stations that together will be used for data entry. The second system is not simply a key-to-tape system but an additional computer with multi-programming capability that will be used for key-to-tape data entry. The system will require programmer support for operation but will provide limited editing capability in return. There is a plan to increase the number of stations to eight at a later time. This keying system should be installed and operational by mid 1981.

The main system will be operated as a time-sharing system with four stations initially. More stations will be added once the planned DOA center is completed and additional terminals may be hard wired to the central system. There is a limit of 24 active stations imposed on the system but there are ports for 32 stations. There is no plan to attempt the use of telecommunications for connecting remote terminals in the offices of the different Directorate General in Jakarta and Bogor in the four or five years that must intervene before the completion of the planned DOA facility in Pasar Minggu. Instead, a courier service will be used to transfer personnel and/or programs, data and output between the center and the user Directorates General.

b. Software

(1) In addition to basic system software, the hardware contractor is obligated to provide the following software:

a COBOL ANSI 74 compiler,

a FORTRAN IV compiler with accompanying FORTRAM library,

a COBOL report writer package,

RPG - II,

IDS 2 (Integrated Data Store) with data base administrator aids,

sort/merge utilities,

an interactive, CRT-oriented text editor, and

a linear programming package.

Although there was no contractual arrangement for its provision, the hardware contractor promised the availability of SPSS (Statistical Package for the Social Sciences), a promise which he now appears unable to fulfill due to the failure of a contract between SPSS and Kansas State University to produce an SPSS package operable on the Level 6 hardware. The lack of a working SPSS package will prove a major setback to the establishment of this data processing center.

(2) A two-year contract in the amount of approximately U.S. $250,000 has been signed between the DOA and P.T. MASRIDA DATA & ELECTRONICS, a new local computer software firm, for the design of an agricultural data base.

The MASRIDA staff is headed by Tan Soei Tien who holds a Ph.D. degree in linear programming from the University of Zurich and David M. Silverman, a consultant who holds a B.Sc. in computer science from New York University. Both men have a broad range of computer experience. The remaining six members of staff, with one exception, appear to have limited computer training or experience.

Under the terms of the contract, P.T. MASRIDA is obliged to design a data base integrating six areas of DOA activity and interest, (1) projects, (2) finance, (3) personnel, (4) production and price monitoring, (5) soils and (6) library resources.

Any one of these six areas could, in itself, prove to be an ambitious undertaking given the fact that the DOA is an organization with 100,000 employees spread across seven seemingly autonomous units. When it is considered that item 4 (production and price monitoring) might be interpreted to mean the design of a data base to include virtually all data routinely processed by the Directorates General of Food Crops, Estate Crops, Fisheries, Forestry and Animal Husbandry the task appears even more formidable.

Unfortunately, the track record for software contracts is a disappointing one. The more diffused the focus of the project the more complex is the design and the greater are the demands placed upon the ultimate user of the system to guide the software contractor in the design. The experience of the Directorate General of Estate Crops, whose attempt to encompass personnel, inventory, finance, resources and estate statistics into a straightforward reporting system in one year left them with perhaps only two satisfactory reporting systems in personnel and inventory, might serve as a warning against spreading the available resources too thin.

If the data base design does undertake to include provision for routinely collected data on crop area, crop yield or crop production, the software contractor is going to have to face directly the issues of what the source of these data will be. As we have already described, there are two quite distinct and independent sets of edited data at the level of greatest resolution, the subdistrict level. Should the data be taken from DOA sources, there will be the obvious disadvantage

that the data subsequently retrieved from the system will differ from officially published statistics. If the CBS should be chosen as the source for the data perhaps it would be best if such an undertaking were simply postponed until the CBS could be involved in it as well.

Under the circumstances, it would appear that this decision ought to be reviewed by the Steering Committee for Agricultural Statistics of BAPPENAS prior to any attempt to design such a data base.

The success of this software contract, as with all others, will be largely dependent on the various DOA officials ability to state, clearly and in considerable detail, the needs which they will bring to the data base system once it is complete.

(3) As a need develops, it may be worthwhile for the center to request the installation of COCENTS, a COBOL coded version of the CENsus Tabulation System. Although COCENTS has not been tested on the Honeywell Level 6 system, the modifications for its successful adaptation are unlikely to be extensive.

Once COCENTS has been installed, training for staff of the Center and others who will use the package could be arranged either through the CBS in Jakarta or the International Statistical Programs Center.

(4) The ultimate installation of SPSS is probably going to be critical to the success of the center and every effort should be made to persuade the hardware supplier to put pressure on the contractor responsible for the SPSS conversion. Should SPSS prove ultimately to be unavailable, some alternative package such as P-Stat, which is written in FORTRAN, might prove sufficiently portable to be of use.

There really is no acceptable alternative to such a package for much of the work that DOA personnel will want to do using the computer facility. Custom-designed statistical programs are notoriously unreliable and one of the strongest arguments for using a statistical package such as SPSS is its reliability; if the proper data is input, one can have a high degree of confidence that the values written out are what the labels say they are. Only thorough testing of a program can give this assurance and that is an ideal seldom realized with programs written hurriedly for a one-time use.

It might also be advisable for the center to consider the purchase of another compiler or interpreter to provide one of the newer and simpler programming languages that are now available. Both BASIC and PASCAL were designed with the training of new programmers in mind and both would make the system more approachable by individual

users such as researchers in AARD. Of these two possibilities, PASCAL would be the better choice because of its more readable format and its greater potential for generalized use.

c. Staff/Training

It is planned that the computer center will have a professional staff of 21 persons in three years time. A detailed program of training for each of the first 14 members of this staff has been set in motion by the Iowa State Contractor.

Four members of the computer team have returned from 10 months training in computer programming at National Computer Institute (NCI) in Manila and a second group of five will complete a similar program of training in October 1981. A third group of five will complete English training in September 1981 and will be available for work in the computer center late in 1982 after completing a similar ten-month computer programming course at NCI. A fourth group of five will attend NCI in 1982. In addition, two persons, Jogas Napitupulu and Suroto Adi are expected to return in 1983 after having completed master's degrees in computer science at Iowa State University and the University of the Philippines, respectively.

Initially, it is planned that the four staff members who have completed NCI training will be assigned to the software contractor for the duration of the contract. Together with the staff of P.T. MASRIDA these four programmers will take part in a month-long training program provided by Honeywell as part of the hardware contract. Four of the five programmers returning late in 1981 will be available for applications work; one is to be an operator and system programmer.

The most serious handicap that this staff will have to contend with is the lack of experienced senior programmers and systems analysts. There is going to be a need for technical assistance in this staff for a number of years, until some members of the staff gain enough experience to take responsibility for the design and direction of programming efforts. One full-time systems analyst is going to be required simply to organize and manage the work of the center. If programming work, other than that directed by the software contractor, is going to go on there will be a need for at least one other part-time systems analyst who can plan for and direct the available programmers.

Since full utilization of the Computer Center depends upon the availability of trained computer personnel within the various Directorates General of DOA, it would be advisable for the Computer Center to provide in-house training opportunities. The need for trainees to master English prior to attending computer courses abroad is one of the major obstacles to such training. The addition

to the Computer Center staff of an experienced senior programmer whose sole responsibility would be the organization and teaching (in Indonesian) of computer classes and workshops, would eliminate the need to train prospective programmers in English.

An in-house training program of this sort would further encourage the development of computer skills in the Directorates General by making possible flexible course formats that would avoid the requirement that personnel be released from their ordinary positions for extended periods of time. The fact that trainees would be able to work on the actual Honeywell system they would ultimately use would be another advantage of in-house training and it is possible that the specificity of such training might reduce the attrition that frequently follows computer training for staff.

The key to such a program of training would be the person selected for the training role. In addition to a comprehensive knowledge of computer topics, this senior staff person should have a strong teaching ability and should be committed to the work of teaching. If the person selected should lack teaching experience he/she should be required to complete a course in teaching methods. It would be ideal if an Indonesian programmer or systems analyst could be found for this work but the most obvious obstacle will be the probable inability of the DOA to offer a salary sufficiently competitive to attract an able person. (See the recommendations made under Section e, Purpose/Organization.) If it were possible to hire such a person for the Center staff, much of the already scheduled training might be accomplished more effectively and at less cost by utilizing the in-house program.

In our interviews we have identified two possible projects that could be undertaken by the four available applications programmers late in 1981 or early in 1982 if funds could be found for technical assistance in directing the work of these programmers.

The first of these possible projects involves BIMAS, which conducted an intensification evaluation survey in 1980. The survey involved the completion of approximately 3200 47-page questionnaires. The data have been collected but have not yet been processed because the CBS has been preoccuppied with the 1980 census. The International Statistical Programs Center could provide the technical assistance and training necessary for BIMAS and AARD Computer Center staff to process this survey. This project would be primarily a training exercise.

If such a course were chosen we would recommend that BIMAS immediately identify one person (possibly two persons) from its statistical staff who shows signs of interest and promise in computer programming, provide English training for this individual (or individuals) and a short-term (3 months) training course in Jakarta on the fundamentals of computer programming and an introduction to COBOL.

A second BIMAS staff person would be identified to work with the survey statistician and should be given training in English.

Towards the end of 1981 an ISPC survey statistician and systems analyst could begin work in the design of a system for processing the collected data. Together with the BIMAS staff the ISPC survey statistician would begin the design of table specifications and would work with both the BIMAS staff person and the four AARD programmers to train them in the design and standardization of table specifications with detailed instructions provided by the ISPC systems analyst. Keying of the questionnaires could begin before the end of the year.

In January, 1982, the programming effort could be begun by the team of five programmers (4 from AARD, 1 from BIMAS) under the direction of the ISPC systems analyst. The system of programs that would be implemented would include a data file generation routine, a valid value edit routine, a consistency edit routine, a questionnaire file listing routine, a routine for updating the questionnaire file and programs to generate the required tables. As part of this program ISPC could install and instruct the five member team in COCENTS before the effort to write programs for the tables is begun. We would estimate that such a project could be completed during the latter half of 1981 and would require 3 person-months assistance from a survey statistician and 1 to 1.5 person-years of assistance from systems analysts/programmers. The cost of this technical assistance would be approximately $105,000 to $150,000.[1] (Refer to Table 4).

If the processing of the BIMAS intensification evaluation survey were successful and funds were available, it would be possible to consider a similar project to provide for annual processing of the BIMAS intensification potential survey.

A second possibility would involve the design and implementation of a forest inventory system for the DG of Forestry. Such a project would again involve technical assistance from an ISPC survey statistician in the design of forms and tables and in the training of personnel from the DG of Forestry and AARD in methods and technical assistance from an ISPC systems analyst in the design and direction of the programming effort.

d. The Computer Facility

The computer center is to be housed in seven rooms at an AARD complex in Pasar Minggu until a new consolidated DOA complex is completed, probably in 1986. The facility appears adequate for the present but will undoubtedly experience a shortage of space

---

[1] U.S. $7,000 per month.

TABLE 4: ESTIMATED KEYING AND STORAGE REQUIREMENTS FOR BIMAS DATA

SURVEY	FREQUENCY	NO. FORMS	NO. BYTES OF DATA (IN MILLIONS OF BYTES)	MAN HOURS REQUIRED FOR KEYING [1]	MAN YEARS OF TECHNICAL ASSIST	
					SURVEY STATISTICIAN	SYSTEMS ANALYST/PROGRA
Intensification Potential	Annual	3,250	8.6	1430	0.25	1.0
Intensification Evaluation	Occasional	3,250	22.9 [2]	3818	0.25	1.5

[1] Based on 12,000 keystrokes/hour, 4 keystrokes/item.

[2] Based on an estimated 150 items per page of questionnaire.

both for storage and personnel before the move to the more permanent facility takes place.

e. Purpose/Organization

As one researcher we interviewed expressed it, sometimes supply must be relied upon to create a demand. It was surely the hope of those who first envisioned the establishment of a computer facility within the DOA that the availability of such a center would lead to increased use of statistical data in agricultural planning, more use of modeling and linear programming and gradually more sophisticated analysis.

It is unusual to see the establishment of such a facility before there are clearly identified users but, in the long term, the realization of such a goal does not appear unreasonable. Undoubtedly the process of finding and training users of the system will be a gradual and sometimes frustratingly slow process.

It seems likely to us that the proposed computer center staff will require the technical assistance of at least one part-time systems analyst responsible solely for programming design will be adequate to meet the reasonable programming needs of the DOA for several years to come. Given the scarcity of computer programming experience, it is best that the few people with experience be centrally located where they can benefit from the experience of other programmers and from a broad exposure to user needs. Initially, the different Directorates General should be allowed to contract the center's programmers' time. Gradually, as the Directorates General programmer requirements grow, one or more may recognize the desirability of establishing its own data processing staff.

There are two potential problems that will face the DOA in the development of a computer programming staff. The first of these is the lack of appropriate computer job descriptions and titles. It will be unsatisfactory after a point to have to fit a programmer or a systems analyst into a statistician's slot. New positions will be needed for keyers, keying supervisors, computer operators, computer programmers and systems analysts.

The second of these potential problems is the lure of private computer contractors who are able to offer higher salaries than civil service requirements will allow the government to offer. One approach to this problem is to find a way to offer more competitive salaries. Some governments, facing a similar situation, have found it possible to establish an independent authority able to pay wages exceeding normal civil service wages, bringing salaries into a range that will reduce the attrition to an acceptable level.

PERSONS CONTACTED

Department of Agriculture (DOA)

1. Office of the Secretary General

   - Mr. Hendro

2. Office of the Directorate General of Estate Crops

   Directorate of Development

   - Mr. Rusdi, Subdirectorate of Surveys and Statistics

   - Mr. Sumarto, Subdirectorate of Evaluation

3. Office of the Directorate General of Fisheries

   Directorate of Planning, Subdirectorate of Statistics

   - Mr. Untardjo, Subdirectorate of Documentation

   - Mr. Lasma Tambunan, Subdirectorate of Capture

4. Office of the Directorate General of Animal Husbandry

   - Dr. Sukartono, Directorate of Planning, Subdirectorate of Data and Statistics

   - Mr. Waidi

   - Mr. Andreas Warkidi

5. Office of the Directorate General of Forestry

   Dr. Piran Wiroatmodjo, M.Sc. Directorate of Program

6. Agency for Agricultural Research and Development

   a. Central Research Institute for Estate Crops

      - Mrs. Sabandiyah Pranowo, Director of Administration

      - Mr. Rachmat, Directorate of Publications and Information

b. Center for Agro-Economic Research

- Dr. Sarjafidin, Director

c. Center for Statistics and Data Processing

- Dr. Subijanto, Director
- Mr. Kaman Nainggolan, MS
- Mr. Ato Suprapto, MS
- Mr. Edi Abdurachman, MD
- Mr. Munawar Hanif
- Mr. Sablin Yusuf
- Mr. Yuanto Ismuyatmono

7. BIMAS

Mr. Dudunu, Director, Program Bureau

Central Bureau of Statistics

. Sam Soeharto, Demographics Division

- Mr. Leo Kasenda, Computer Center
- Mr. Yuwono, Computer Center
- Mr. Suwandi, Agricultural and Industrial Statistics

Iowa State Univeristy

- Dr. Gary Vocke, Program Director
- Mr. Daniel E. Peterson, Systems Analyst

USAID/Indonesia

- Mr. Jim Gingerich

Institute Pertanian Bogor

- Dr. R.R. Bernsten

Dept. of Agriculture Food Statistics Office in Karawang, West Java

- Mokh. Ambari Wiryasoedarma, Director

ntral Bureau of Statistics Office in Karawang, West Java

- Aji Satiadji