

Reprinted from

**Symposium on  
Machine Processing of  
Remotely Sensed Data**

**June 29 - July 1, 1976**

The Laboratory for Applications of  
Remote Sensing

Purdue University  
West Lafayette  
Indiana

IEEE Catalog No.  
76CH1103-1 MPRSD

Copyright © 1976 IEEE  
The Institute of Electrical and Electronics Engineers, Inc.

Copyright © 2004 IEEE. This material is provided with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the products or services of the Purdue Research Foundation/University. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to [pubs-permissions@ieee.org](mailto:pubs-permissions@ieee.org).

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

## CURRENT METHODS AND POLICIES OF THE STATISTICAL REPORTING SERVICE

Charles E. Caudill, Director, Research Division, Statistical Reporting Service,  
U.S. Department of Agriculture, Washington, D. C. 20250

### I. ABSTRACT

The Statistical Reporting Service as the primary fact-collecting and reporting organization of the U.S. Department of Agriculture is responsible for the national program of crop and livestock estimates. Over time, methodology underlying this program has evolved to the point where its foundation is firmly based on probability sampling, employing both area and list sampling frames. Basic data collected for this program is based on voluntary reporting under a pledge of confidentiality. This places strict limits and restraints on SRS in terms of sharing with other agencies or groups the basic data or even unpublished estimates used or generated in the crop and livestock estimates program. Disclosure of individual data is prohibited by law or by Department regulation.

### II. INTRODUCTION

The basic job of the Statistical Reporting Service (SRS), of the U.S. Department of Agriculture is to collect and disseminate current data on U.S. agriculture. SRS and its predecessor organizations have been performing this task for more than a century. Today SRS serves as the chief fact-gathering agency of the Department. In addition to the agency's program of crop and livestock estimates, SRS has responsibility for statistical research and service for all of USDA. This function includes review, clearance, coordination, and improvement of statistics for the Department. In addition, as the capability for designing and implementing modern probability surveys has developed, SRS has increasingly been called on to perform statistical services for other Federal, State and private agencies on a reimbursable or advance payment basis. These services include survey design, data collection and processing, and training activities for agricultural data programs as well as providing consultation and technical assistance for other countries.

#### SCOPE OF CURRENT ESTIMATES PROGRAM <sup>1</sup>

The current program of crop and livestock estimates includes some 550 national statistical

reports published each year, plus additional material issued by our 44 field offices. These reports are prepared and published in accordance with a strict timetable which is prepared and announced well in advance of actual estimates. Surveys are conducted during a relatively short time frame. Typically, less than three weeks will be involved from the beginning of data collection through publication for most SRS reports.

The SRS organization includes, in addition to the Washington, D.C. office, 44 State Statistical Offices serving all 50 States. Staffing plans are for 1,139 full-time employees supplemented by up to 1,500 part-time enumerators. In addition, most State offices will include a limited number of State employees, funded through cooperative agreements with State agencies such as State Departments of Agriculture. The organizational structure of SRS is particularly well suited for centralized planning and policy development and for decentralized execution of data collection and processing, and avoidance of duplicated efforts between Federal and State or local agencies.

By law or Department regulation the SRS program of crop and livestock estimates requires National and State estimates. Through the cooperative agreements with State agencies many State Statistical offices also prepare estimates for counties or estimates for other local areas. In addition to the 550 national reports released by SRS, our State offices release some 10,000 reports which are pertinent to individual state agricultural activities.

Reports for more than 120 commodities (including field and seed crops, vegetables, and fruit and tree nuts) provide estimates of acreages which farmers intend to plant; acreages actually planted and harvested; yield; production and crop disposition; as well as periodic indications of remaining stocks for important storable crops. Monthly forecasts of production are prepared for major crops throughout the growing season.

Livestock, dairy and poultry reports cover inventory numbers, births, cattle and sheep on feed, wool and mohair production, slaughter, meat output, milk and egg production, disposition, and income, as well as a host of other related data required to provide timely information on current and potential supplies and production for this important segment of U.S. agriculture.

Price reports include prices received by farmers for about 200 products, prices paid for approximately 600 items farmers buy for production needs or family living, as well as monthly indexes of prices received, prices paid, and parity prices. Also estimates of farm labor and wage rates are published quarterly.

#### METHODOLOGY

Information and data used by SRS are collected by a variety of methods --- mail, telephone, personal interviews and actual field observations. Contacts include farmers, ranchers, hatchery and grain elevator operators, feed and fertilizer manufacturers and dealers, farm supply and equipment dealers, slaughter and other processing plant managers, and many other agricultural experts and firms serving U.S. agriculture.

Nearly all the surveys conducted by SRS are sample surveys based on relatively small samples. Historically, when data collection was almost exclusively by mail, SRS surveys employed nonprobability sampling due to inadequacy of existing sampling frames and lack of resources for followup of nonrespondents. Analytical techniques, using quinquennial Census of Agriculture data as a "benchmark", were developed to remove persistent biases from nonprobability survey results. These procedures were used for many years and provided reasonably accurate current agricultural estimates. However, this system has been largely outmoded during the last decade as implementation of a national probability sample has been completed. The 1964 Census of Agriculture was the last used by SRS as National "benchmark data" for any of the major crop and livestock items included in the program of current statistics.

#### Probability Sampling

The first requirement for probability sampling is an adequate sampling frame.<sup>2</sup> As methods employed by SRS evolved from a nonprobability system of surveys to one based primarily on probability sampling, the main problem was the acquisition of an adequate sampling frame. When this transition began in the early 1950's, the most obvious choice seemed to be a List Frame, particularly since SRS already used many lists containing names and addresses of farm operators for the existing nonprobability mail surveys. However, after some investigation it was determined that the lists then available to the agency, singly or in combination, were completely inadequate as sampling frames for probability sampling. Also, in the time before the advent of the computer it was almost physically impossible to maintain and manage a large list of names.

The second frame studied by the agency was the "Master Sample Area Frame" constructed by the Bureau of Agricultural Economics and the Bureau of the Census at Iowa State University in the early 1940's. This frame proved satisfactory for the Corn Belt States and most of the Southern and Eastern States, so implementation of a probability area sample was begun in the mid-1950's on a State-by-State basis in these areas of the U.S. The designers of the Master Sample materials recognized that the stratification (open country, urban places, and rural places) was inadequate, particularly in the Western States. The need for a new area sampling frame was still evident in the late 1950's.<sup>3</sup> Therefore, in 1960 SRS undertook the task of constructing a new area sampling frame in 11 Western States. More detailed land use strata were defined during the mapping process which was completed by 1964 in the West and the Northeastern States and Florida by 1965. Since then most of the Master Sample Area Frame has been replaced by new land use area frames constructed by SRS or for SRS by Iowa State University. Once this task is completed for all States, plans call for updating the SRS frame with recent changes that have occurred in agriculture and land use patterns. This updating work should begin in late 1977 for California, with other states to follow as rapidly as resources will permit. It is anticipated that LANDSAT imagery will be quite helpful in bringing land use stratum boundaries up to date.

By 1965 the task of implementing a national probability general purpose survey was completed with a sample of some 17,000 area segments (sampling units selected from the area frame) enumerated during the May-June period each year. This survey, generally known as the June Enumerative Survey (JES), provides a probability foundation for the crop and livestock estimates program, particularly for major crop acreage estimates. It also provides a base for probability selection of fields for use in objective yield surveys, again for major crops such as wheat, corn, cotton and soybeans.

#### June Enumerative Survey

Because of its importance as the major SRS probability survey for collecting crop acreage data, I have included the following brief description of sample design and survey procedures for the JES.

The current sample for the JES consists of approximately 16,000 area segments (sampling units, typically about one square mile in size, although they may vary depending on intensity of agriculture). This sample is designed as a stratified sample with states and areas within states serving as strata. For states where the Master Sample frame is used, crop reporting districts serve as intrastate strata. Crop reporting districts are groupings of contiguous counties having relatively similar agricultural characteristics. Typically, States contain about nine districts. In States using the SRS land use frame, the sample is allocated to the land use strata within a state or even within a smaller geographic area. The sample allocation provides national and reasonably precise state estimates.

Sample selection within strata has generally followed a systematic approach using a geographically arrayed listing of the frame. Recently, as new samples have been selected from the SRS land use frame, interpenetrating sample designs have been used. This design uses several smaller independent within stratum samples. Substantial gains in efficiency have been realized through this change in terms of increased precision of JES estimates, and ease of selecting segments for the original sample and for the sample rotation scheme. Typically, 20 percent of the JES segments are rotated annually, primarily to reduce respondent fatigue.

Data collection for the JES involves an extensive training program before each survey. Each year the Washington staff conducts regional schools to train survey statisticians from our state offices. State survey statisticians then hold schools within their States for training enumerators (or interviewers) who will visit each area segment to obtain the required information. In addition to the careful selection and training of enumerators, a number of other quality controls are included in JES operating procedures. These include detailed instruction manuals and materials, close field supervision of enumerators, built-in questionnaire checks, and reenumeration of a subsample of segments by supervisors. We also use aerial photography to locate and identify the segments and compare acreages reported by farmers with measured acreage on the photographs.

Enumerators use State and county road maps to locate segments and aerial photos to define precisely the segment boundaries. The enumerator will interview each person who operates land within the segment to obtain information about acreage and land use for every field inside the segment. This information, which includes both net and gross acreage for fields involving planted crops, is carefully recorded on the questionnaire for later summarization in the State office.

It should be noted that the JES collects information on livestock, farm labor, stored grains, population, and other economic and social characteristics pertaining to farms or farm operators. In fact, only about 20 percent of a typical JES questionnaire will be devoted to crop acreage questions. This means that even if remote sensing could entirely replace the present data collection procedure for acreage only very minimal savings would accrue and these would probably be more than offset by the higher data processing costs for classification and summarization of the remotely sensed data.

#### Multiple Frame Sampling

Between 1955 and 1965 installation of a national area sample was the highest priority effort in SRS. During this period the new probability area samples supplemented the older nonprobability list samples, and use of lists was not abandoned. Two developments during this decade served to increase interest in building and using lists for probability sampling. By 1965 ADP

technology, both hardware and software, had evolved to the point where handling large lists was no longer the formidable task it seemed in the early 1950's. This technology permitted the easier resolution of many list inadequacies, such as duplication. Also, in 1962 at the Minneapolis meetings of the American Statistical Association Dr. H. O. Hartley presented a paper titled "Multiple Frame Surveys," which provided a relatively simple theoretical base for the joint use of more than one sampling frame for probability sampling. Specifically, this led to the development of procedures for using incomplete list frames where data collection by mail is relatively inexpensive, in combination with the more expensive but complete area frame. Much research and effort has been devoted to multiple frame surveys by SRS during the past decade. While most of this work was aimed at improving the precision of livestock estimates the procedure has also been used for crop acreage estimates in states having appropriate list frames. Beginning during the current fiscal year (FY 76) a major effort has been mounted by SRS to provide a reasonably complete and consistent general purpose list frame by the end of 1977 for all States except Hawaii and Alaska. This frame, with appropriate control information for stratification, and provisions for continuous updating, will permit use of the multiple frame technique for new probability surveys with a minimum increase in cost.

#### Objective Yield Surveys

While the area and multiple frame surveys generate very precise estimates of crop acreages planted and harvested, other methods are used by SRS to forecast and estimate crop yields. Objective yield surveys provide crop yield information for estimates (at harvest) or forecasts (prior to crop maturity) based directly on counts, measurements, and weights of the crop made from small randomly selected plots in a probability selection of sample fields. Prior to maturity, plant counts and observations correlated with final yield are used to make yield forecasts. Once the crop is mature, the sample plots are harvested, weighed and expanded to a biological yield per acre. Enumerators glean of special sample plots after field harvest by the grower to provide a measure of harvesting loss for adjusting the biological yield to arrive at a net harvested yield. This net yield multiplied by harvested acres provide an estimate of production. Objective yield surveys for wheat, cotton, corn, soybeans, potatoes, and citrus are used by SRS to provide probability estimates of yield. These surveys are supplemented by extensive mail surveys of producers, particularly just after harvest to obtain additional yield and production data for these and other crops.

The current SRS system of probability surveys provides independent estimates with known measures of precision (sampling errors). Typical sampling errors are shown for a few major items in the following table.

Table 1. Coefficients of Variation, June Enumerative Survey and Objective Yield Survey, U.S., 1975

| Item          | Coefficient of variation |       |            |
|---------------|--------------------------|-------|------------|
|               | Acres planted            | Yield | Production |
|               | ----- Percent -----      |       |            |
| Winter wheat  | 1.5                      | 1.0   | 2.0        |
| Corn          | 1.1                      | 0.9   | 1.7        |
| Cotton        | 3.5                      | 1.0   | 3.7        |
| Soybeans      | 1.7                      | 2.1   | 2.1        |
| Grain sorghum | 3.3                      | -     | -          |

SRS POLICY REGARDING USE OF UNPUBLISHED DATA AND SAMPLING FRAME MATERIALS

As research scientists, particularly within the remote sensing community, have become increasingly aware of the SRS area sampling frame and the primary data collected through the JES and other surveys, the number of requests for sharing these materials and data has risen substantially. While SRS has been able to satisfy some of these requests, many cannot be met for technical, legal, administrative, or other reasons. Therefore, to acquaint the scientific community with SRS policies for sharing such things as sampling materials, unpublished estimates, and sample data, I will briefly outline our procedures.

Individually Reported Data

Practically all primary data used in the crop and livestock estimating program is provided on a voluntary basis by respondents to SRS surveys. This dictates that the data be collected under a pledge by SRS to maintain the data as confidential and to avoid publishing or revealing to anyone outside the agency any information which would violate this pledge. A violation of this pledge would "dry-up" future voluntary reporting. Disclosure is specifically prohibited by law ... 18 U.S.C., Section 1905. Under "disclosure of confidential information generally," the law states "Whoever, being an officer or employee of the United States or any department or agency thereof, publishes, divulges, discloses, or makes known in any manner or to any extent not authorized by law any information ... confidential statistical data ... shall be fined not more than \$1,000, or imprisoned not more than one year, or both; and shall be removed from office or employment." We interpret this restriction as applying to the identity of the respondent as well as to the individually reported data. Thus, the SRS lists of reporters are protected from those who seek them for regulatory, political, or commercial interests. Similar restrictions are applied to lists of individuals and

to individually reported data by the Bureau of the Census in the Department of Commerce. Also, in our publication of statistical data, any data cell which represents less than three respondents is suppressed to avoid disclosure of individually reported data.

All of the data collected through the JES and objective yield surveys as well as other surveys is treated as confidential and cannot be shared outside SRS. In certain instances where all identifiers, names and addresses, and segment location, have been removed it is possible to provide individual sampling unit data for legitimate research requirements. Exact sampling unit location cannot be divulged, since doing so would provide others with a random sample which could be used for other surveys and could create such a respondent burden that it would have a very serious effect on voluntary response to future SRS surveys.

Based on the preceding legal and other considerations SRS policy for use of unpublished sample data can be outlined as follows:

- (a) use permitted only after personal (including location) identifiers have been removed,
- (b) must be used only for the public benefit,
- (c) use permitted only for objective studies in methodology, by recognized government research agencies, universities, and private research organizations whose purpose is devoted to the public interest,
- (d) use permitted only after the data are no longer current (after the end of the year or crop season to which the data are applicable),
- (e) data will be provided only under a formal written agreement between SRS and the requesting agency showing intended use of the data, designating by name, the individual who will be responsible for use of the data, and giving SRS the right to review, for statistical standards, reports on such studies before publication.

Unpublished Aggregated Data

SRS in the course of preparing state and national estimates summarizes most survey data at additional levels. These data may include stratum summaries, either for geographic (crop reporting district, county, etc.) or for size group strata within states. These data have limited or uncertain reliability but can generally be provided for uses which are clearly in the public interest provided the user understands that because of the lack of precision the data must be used with extreme caution. Also the user must agree to SRS review of any publications derived from these data, for statistical standards, and to determine that the manner in which they are used will not be detrimental to the voluntary reporting structure under which SRS operates. Although these kinds of records have been identified under the Freedom of

Information Act as exempt from public disclosure,<sup>4</sup> it is SRS policy to attempt to meet reasonable requests to the extent they can be provided without revealing individual operations. In some cases, administrative restraints make it impossible to provide certain data. For example, because of the large volume of data collected each year by SRS a fairly rigid disposition schedule for destroying individual questionnaires and records must be followed. Therefore, these records are maintained only as long as absolutely necessary, usually not more than one year and sometimes less. Field and tract boundaries on aerial photos used in the JES are erased prior to each survey because the same photos must be used for more than one survey. This means that historic JES data cannot be related to tract and field boundaries, since the boundaries may change from year to year.

#### REFERENCES

1. For more detail refer to "Scope and Methods of the Statistical Reporting Service," Miscellaneous Publication No. 1308, USDA, SRS, July 1975.
2. B. W. Kelly, "Probability Sampling in Collecting Farm Data," J. Farm Econ. 45: 1516 Dec. 1963.
3. Harold F. Huddleston, "A New Area Sampling Frame and Its Uses," J. Farm Econ. Vol. 47. No. 5, Dec. 1965.
4. Federal Register, Vol. 32, No. 128, Tuesday, July 4, 1967, page 9728.

#### Sampling Frame Materials

The area sampling frames constructed and/or used by SRS represent an important capital asset and as such should be used as widely as possible for public benefit. SRS encourages this use by:

- (a) Selecting special area samples for data collection under the SRS statistical service function. Data collection is usually performed by SRS but could be performed by a properly qualified and authorized agency for whom the sample would be selected.
- (b) Allowing other agencies or groups to use the materials for selecting their own samples. This activity must be accomplished under SRS supervision at the site where the sampling frame materials are housed since the materials are used on a day-to-day basis for SRS survey activities.

The above activities must be closely coordinated with ongoing SRS activities in order to avoid disruption of normal work and to assure that use of these materials is for public benefit.

SRS does not presently have a comprehensive general purpose list sampling frame for agriculture, but work is underway to construct such a frame on a State-by-State basis. Control of this frame will be decentralized to each SRS state office and will be maintained at this level. Since this list will contain control information used for sampling, much of which will be collected through SRS survey activities, its content will not be revealed except for individual records to individuals included on the list as required under the Privacy Act of 1974. Public Law 93-579, the Privacy Act permits individuals to have access to records pertaining to themselves. Use of this frame can be shared through the SRS statistical service function when SRS serves as the contractor, conducting a complete survey, including sample selection as well as data collection and perhaps data processing.