The New California Area Frame

A Statistical Study

Ron Fecso
Van Johnson

THE NEW CALIFORNIA AREA FRAME: A Statistical Study, by Ron Fecso and Van Johnson. Statistical Research Division; Statistical Reporting Service; U.S. Department of Agriculture. SRS-22.

ABSTRACT

The use of new technologies, such as LANDSAT imagery and updated mapping materials, has led to increased accuracy in the California area frame (the frame delineates all California and serves as the basis for statistical sampling). However, the recently developed area frame analysis package uncovered some minor sampling and nonsampling errors in the 1979 June Enumerative Survey, which limited the gain in efficiency. Alternative procedures for stratification, sample allocation, and improvements in quality control were developed during this study to aid in the reduction of future survey errors.

Keywords: Area frame, frame development, sample allocation, sample size, statistical design, survey sampling.

ACKNOWLEDGMENTS

We thank the members of the Sample Survey Research Branch and those of the Methods Staff for comments and editorial assistance of the preliminary draft of this report. We are especially grateful to Martha Farrar, Barry Ford, Jack Nealon, Jim Carlin and Chap Gleason for their help with the organization and content of the very lengthy preliminary draft. The California State Statistical Office provided data and the Fairfax Unit and Systems Branch provided technical assistance in obtaining the data tapes. Jennifer Moerman, Karen Watkins, Phyllis Piper, Sue Dryer, and Barbara Manthos typed the revisions to this manuscript. We appreciate the management support of Norm Beller, Bob Tortora, and Wayne Gardner during the development period of the area frame analysis package.

Washington, D.C. 20250 September 1981

CONTENTS

GLOSSARY

Ancillary data--Any information that is supplemental to the required data for constructing an area frame. This information can include estimates from previous surveys and county estimates.

Area frame--The aggregation of all parcels of land for a specific geographic locale.

Color signatures--The colored representation of cropping patterns in the area scanned. For example, bare soil might have a black signature; crops might be bright red, exhibiting regular patterns (fields).

Cultivated land--Considered to be agricultural land excluding noncropland uses. The value is computed by subtracting the acreages for pasture, farmstead, grazing land, wasteland, and idle land from the total agricultural land acreage.

Digitizing--Electronically computing a digitized area from a series of points defining the boundary.

The June Enumerative Survey (JES)--The JES is conducted annually in the 48 contiguous States during the last week of May and first week of June. Data collected for SRS reports concern crop acres and land use, livestock inventories and births, farm labor and other economic factors, farm population, and number of farms. The basic area frame sample used by SRS includes about 15,000 area segments. The number varies by State according to land area, importance and diversity of agriculture, and age and precision of the area frame. Most midwestern States have about 350 segments and southern States have about 450. Texas and California have the most segments, with 848 and 981, respectively. The area segments are completely enumerated; that is, all land within the sample must be accounted for and identified as to use. The segments include about 100,000 separate tracts, each represented by a different operator, who is contacted in person for information. About half of these tracts have agricultural activities.

LANDSAT imagery--The LANDSAT imagery used in developing the California area frame is derived from data collected by the multispectral scanner (MSS) aboard the LANDSAT satellite. The MSS picks up reflected and emitted energy of a scene in a line-by-line fashion. The optics of the MSS system refract this beam of energy, separating it into four components--green, red, and two near-infrared components. The response in each wavelength band is then stored on magnetic tape in digital form. It is from the MSS digital data that MSS image products are created. In the false-color MSS imagery which we use, three primary colors are assigned to three of the four LANDSAT bands. The varying levels of the MSS receive energy in these bands and mix the primary colors in varying proportions, thus creating a colored representation of the area scanned by the LANDSAT satellite.

Large farms--An operation is a large farm if its holdings are so vast and production so diversified that data collection on a tract basis is virtually impossible. The operations on the large farm list are not sampled but are included in the JES with a probability of one. Since the large farms are completely enumerated, any data for tracts in the area frame which are part of a large farm are excluded from the expansions. In essence, the area estimate is a multiple frame estimate which combines the segment expansions and large farm enumeration to derive the State estimate.

Photo index--A composite of individual frames of low-level aerial photography which depicts the land area for a specified location and provides a reference system for ordering enlarged photos of smaller areas.

Planimeter--Measuring an area manually by following the stratum boundary with an area measuring instrument.

Segment--The area frame sampling unit. That is, a piece of land with boundaries which can be delineated on a map such that each count unit is composed of a defined number of nonoveralpping segments. All parcels of land in count units are contained in these segments.

Stratifier--A person who classifies the land area into the various land use stratum. These individuals interpret the LANDSAT scenes as well as the low-level photography and draw the stratum boundaries on the framework.

Target segment size--The sample unit size that is desired for each stratum. All segments drawn in a count unit are expected to be within a given tolerance of the target segment size.

Tract--The reporting unit for area frame surveys. Since area frame sampling is a form of cluster sampling, several tracts may fall within each segment.

The New California Area Frame
A Statistical Study

Ron Fecso and Van Johnson

INTRODUCTION

Area frame surveys are a major source of indications for U.S. agricultural estimates. The efficiency of area frame surveys used by the Statistical Reporting Service (SRS) is improved through the use of a stratified design based on land use. The efficiency, however, may diminish over time as land use changes and eventually makes it necessary to construct a new area frame.

A new area frame was developed for California and first used in the 1979 June Enumerative Survey (JES). 1/

Analyses utilizing the Area Frame Analysis Package (AFAP) show that the new frame for California is more efficient than the old frame. 2/ The replicated design allows for easy reallocation of the sample whenever required in order to provide reduced sampling errors or to alter survey costs. Reliable estimates for crops previously having coefficients of variation (CV's) outside the usable range can be achieved through reallocation and changes in the frame construction procedures. 3/ The use of specialty strata improves the efficiency of the stratified design.

Some problems were observed in the stratification, segment size distributions, and segment content which increased sampling errors. Errors in mapping, enumeration, and summarization led to minor nonsampling errors; however, improved procedures for quality control should lessen their impact on future survey results.

The results of this analysis of 1979 June Enumerative Survey data provided both detailed information on the success of

---

1/ Terms that are defined in the glossary are underlined the first time they appear. Terms that may be unfamiliar to many readers will be defined in footnotes.

2/ The Area Frame Analysis Package was developed by the authors during 1979/80. A detailed description of the analysis package will appear in a future publication.

3/ For the JES, the CV's for major crops are expected to be in the 2- to 4-percent range at the national level and 3 to 12 percent at the State level.

the new frame and pointed out areas where improvements in procedures for constructing area frames can be made.

FRAME DEVELOPMENT

The Sampling Frame Development Section (SFDS) is responsible for constructing and maintaining SRS area frames. Individual frames, which are aggregations of all parcels of land, have been developed for each State. Thus, the sampling design can be tailored to the land use and agricultural practices in each State.

This section contains a description of the sampling design and frame construction procedures used in California. 4/ The California frame is unique since LANDSAT imagery was used for the first time in the construction of an operational frame (9). 5/

Frame Design

Early experience with area frame surveys indicated that stratification according to land use was essential (8). The area frames used by the Statistical Reporting Service (SRS) are usually stratified into six general land use strata based on the amount of land cultivated, and can be further subdivided based on varying percentages cultivated. The general land use strata are intensive agriculture, extensive agriculture, cities and towns, range, nonagriculture, and water.

The new California frame, however, was not limited to these six land use strata. Instead, information obtained from ancillary data was used by the research group of SFDS to develop crop-specific strata within the general land use framework.

The strata definitions for the new California area frame are:

Stratum 13--Fifty percent or more cultivated, mostly general crops with less than 10 percent fruit or vegetables.
Stratum 17--Fifty percent or more cultivated, mostly fruit, tree nuts, or grapes mixed with general crops.
Stratum 19--Fifty percent or more cultivated, mostly vegetables mixed with general crops.
Stratum 20--Fifteen to fifty percent cultivated, extensive cropland and hay.

4/ Much of the general description on frame construction is drawn from Houseman (5). A pilot study on the use of LANDSAT in area frame construction was conducted in 1977. A description of the pilot study is found in (4).
5/ Underscored numbers in parentheses refer to literature cited at the end of this report.

Stratum 31--Agri-urban, more than 20 dwellings per square mile, residential mixed with agriculture.
Stratum 32--City, more than 20 dwellings per square mile, heavily residential/commercial, virtually no agriculture.
Stratum 41--Privately owned range, less than 15 percent cultivated.
Stratum 43--Desert range, barren areas with less than 15 percent cultivated, virtually no crops or livestock.
Stratum 44--Public grazing lands, Bureau of Land Management or Forest Service grazing allotments.
Stratum 45--Public land not in grazing.
Stratum 50--Nonagricultural, includes State and National Forest, wildlife refuges, military reservations, and similarly designated land.
Stratum 62--Known water (not sampled), larger than one square mile in area.

Frame Construction

The construction of a Statistical Reporting Service area sampling frame requires a major investment. An efficient frame requires precise stratification and a design which will not become out of date in a few years due to land use changes. To increase the long-term efficiency of area frames, land use stratification and count units are used to obtain economies in sample design, sample selection, and frame maintenance. 6/

The count unit is an area of land larger than a segment (area frame sampling unit) but smaller than the smallest political subdivision. Several factors had to be considered in determining the specifications for the count units for the California frame. These factors included:

(1) The availability of suitable boundaries--count unit boundaries should be easily recognizable and have a high degree of permanency.

(2) The determination of count unit size--count units should be large enough to allow for alternative specifications of segment size.

(3) The economy in the selection of area samples--only those count units that contain sample segments need to be

---

6/ Count units in early area frames were a count of farms indicated on highway maps. Houseman (5) has suggested discarding this term for a more general one such as "frame unit" but count unit is still used by SRS.

broken down into segments, thus reducing the number of boundaries drawn on the framework and the cost of area frame sampling. The greatest economies occur when the number of count units is much larger than the number of segments to be sampled, and the number of count units is much smaller than the number of sample units in the frame.

(4) The availability of ancillary information--several sources of data are incorporated into the construction of count units. These sources included LANDSAT imagery, prior survey data, census data, low level aerial photography, and knowledge of cultural practices.

(5) The definition of the population--the survey population and subpopulations often are defined not only in terms of reporting units but also in terms of geographic coverage.

(6) The approach used to set the specifications--although the specifications could be set using size and topographic landmarks primarily for the boundaries, the SFDS used the ancillary data such as land use to achieve homogeneity within the count units.

Land use strata can be constructed in two ways: count units can be formed by using ancillary data to classify parcels of land, or alternatively, count units can be constructed based on land use as determined by aerial photography, or satellite imagery. SFDS uses the latter in constructing area frames for SRS surveys. This method allows:

(1) The frame to be easily updated by reclassifying the count unit to the stratum which reflects the land use changes,

(2) Stratification to achieve homogeneity within a stratum,

(3) Count unit boundaries to coincide with areas that might be used as domains of study, and

(4) Count units to be formed according to the best size or boundary criteria for each land use stratum.

Classifying all land into one of the defined strata was the first step in constructing the California frame. Overlays of the county maps for California matched the same scale as the LANDSAT imagery, and each county was located on the LANDSAT scenes. 7/ Overlays showing the locations and crop contents of segments sampled during prior surveys were then oriented on the scene. A stratifier used the data collected during prior surveys, county estimates, crop calendars, general background information such as field sizes and the date of the LANDSAT image, and other available ancillary data to interpret the color signatures on the LANDSAT scenes. For example, bright red areas in regular patterns might indicate crops, irregular patterns might be native vegetation (range), and bare soil signatures in regular patterns might indicate land under cultivation.

The stratifier was able to break down each work unit into the land use strata defined for California by interpreting the LANDSAT imagery and extrapolating the prior JES and other ancillary data. Strata boundaries were then delineated on the framework. 8/ Photo indexes (PI) helped identify permanent boundaries not readily apparent on the county highway maps. After the strata were set for the work units, each stratum was broken down into count units. The size of the count unit was measured by digitizing the area. The number of segments in a count unit was then determined by dividing the count unit area by the target segment size. After the number of segments was assigned to each count unit, the count units in each stratum were arranged by county to group count units which were agriculturally similar. 9/

To develop the geographic stratification, the arranged count units were listed with the associated number of segments in each count unit and the total number of segments accumulated for the stratum. Paper stratification is the form of geographic stratification used in the SRS design. The paper strata are formed from sequential groups of segments as defined by a count unit ordering. This ordering is such that the groupings are of equal size within rounding and the content of the groups is agriculturally similar. Thus, the geographic

---

7/ The total workload was divided into work units, each consisting of a single county, which kept workloads for individuals at manageable levels and allowed an even flow of work through completion of the frame.

8/ The framework is a complete set of county maps which are maintained as the official documentation of the frame.

9/ This ordering is essential to the SRS geographical stratification. Originally factor analysis on county data was used to develop the county ordering (7). The current procedure of using cluster analysis on scaled county data was developed by Fecso and is available in a working paper.

proximity and agricultural similarity of these groups of segments results in the paper strata being almost a crop-specific stratification.

Within a land use stratum, simple random samples of equal size are drawn in each paper stratum. Sequential samples drawn in each paper stratum are assigned to the sequential replicates of the land use stratum. Thus, the design within a land use stratum might best be described as a replication of drawing a single random sample from each group of a systematic grouping of segments. The design not only contains the desirable properties of systematic and replicated sampling, but as shown later in the paper, provides for more flexibility in sample allocation and design changes.

The random selection of the segment was a two-stage process in which the count unit containing the segment was first selected with probability proportional to the number of segments in the count unit. The selected count units were then located on the PI's and divided into the assigned number of segments. Generally a "large" count unit (15 or more segments) was split into two or more parts of about the same number of segments. One of the split portions was selected with probability proportional to the number of segments, and the selected portion was divided into segments. After the count unit or split portion had been broken down into segments, one of these segments was selected at random and coded for identification. The photo indexes were then given to cartographers, who copied the segment boundaries onto an enlarged photograph to be used in field enumeration.

All strata were constructed and sampled in the above manner with the exception of stratum 44, the public grazing lands stratum. This stratum was not broken down into count units. Instead the sample units were boundaries of the grazing allotments, causing segment size to vary widely. In 1979, the sample segments were selected with probability proportional to land area, but in future surveys the sample units may be selected with probability proportional to the number of cattle permitted on the grazing allotment.

FRAME COMPARISONS

To compare the efficiency of the new frame, we obtained the data for the new frame from the 1979 JES while old-frame data came from the 1978 JES survey. Exact stratum-by-stratum camparisons are not possible because strata definitions were changed from the old to the new frame. In the frame that was constructed during the 1963-64 period, four general land use strata were broken down into more specific land use categories. 10/ The 1978 JES Supervising and Editing Manual

10/ The old frame was updated during the period it was in use.

gave the following outline for the old frame used in California. The comparable land use strata for the new frame is indicated in the parentheses.

1. Cultivated land (13, 17, 19, 20)
   - Dryland
   - Irrigated

2. Cities and towns (31, 32)
   - Population less than 7,500
   - Population 7,500 to 49,999
   - Population 50,000 or more nonindustrial
   - Population 50,000 or more industrial

3. Nonagricultural (50, 62)

4. Range (grazing) (41, 43, 44, 45)
   - Grazing allotments administered by the Forest Service, Bureau of Indian Affairs, or other governmental agency
   - Privately owned grazing land plus grazing allotments administered by the Bureau of Land Management
   - Lands with little suitability for grazing such as rough terrain, timber, and arid lands

The Enumerative Summary System of SRS computed all of the estimates and sampling errors for 1978, while those for 1979 were computed by both the Enumerative Summary System and the AFAP summary module. AFAP data were corrected to reduce some of the nonsampling error, but the summary values were biased slightly downward because the observed values (recorded to tenths in the JES master record) were truncated to whole numbers in creating the data set. However, the bias due to truncation had minimal impact on the AFAP expansions and their standard errors. 11/

Table 1 shows the land use strata for the four general categories as defined for the old frame. The values for the general land uses were obtained by aggregating the individual land use estimates and computing the standard error for this variable within each land use stratum. The frame base acres are the sum of count unit sizes, as measured from the framework. Count units in the old frame were planimetered while those in the new frame were digitized. The frame base acres for both old and new frames are within the acceptable tolerance level of ±1

11/ The truncation problem has been corrected in a later version of AFAP. Since the impact on the California results was minimal, we did not feel that we needed to recreate the data set for this analysis.

Table 1--Frame comparisons for California, 1978-79

General land use strata	Frame base 1/	Enumerated area 2/	Estimate	Standard error
		Acres		
1978:				
Cultivated land	13,853,440	530,473	14,239,755	167,756
Cities and towns	1,568,640	1,083	1,791,757	187,023
Nonagricultural	2,967,680	245,184	2,967,765	11
Range	81,879,040	NA	72,173,937	6,738,423
Total 3/	100,268,800	NA	91,173,214	6,743,104
1979:				
Cultivated land	15,031,680	350,280	15,019,541	108,619
Cities and towns	3,843,200	7,204	3,998,707	142,135
Nonagricultural	9,069,440	10,076	8,924,678	175,134
Range	71,669,760	NA	69,510,342	561,783
Total 3/	99,614,080	NA	97,453,268	615,038

NA = Not applicable because of design changes between years.
1/ Planimetered for 1978, digitized for 1979.
2/ Excludes large farm acreage which is completely enumerated.
3/ Excludes water area greater than 1 square mile.

percent of the census land area (100,069,000 acres). Enumerated area is the land area from which data were collected during sampling. The estimates and standard errors were computed from the standard formulas for the stratified design used in the survey.

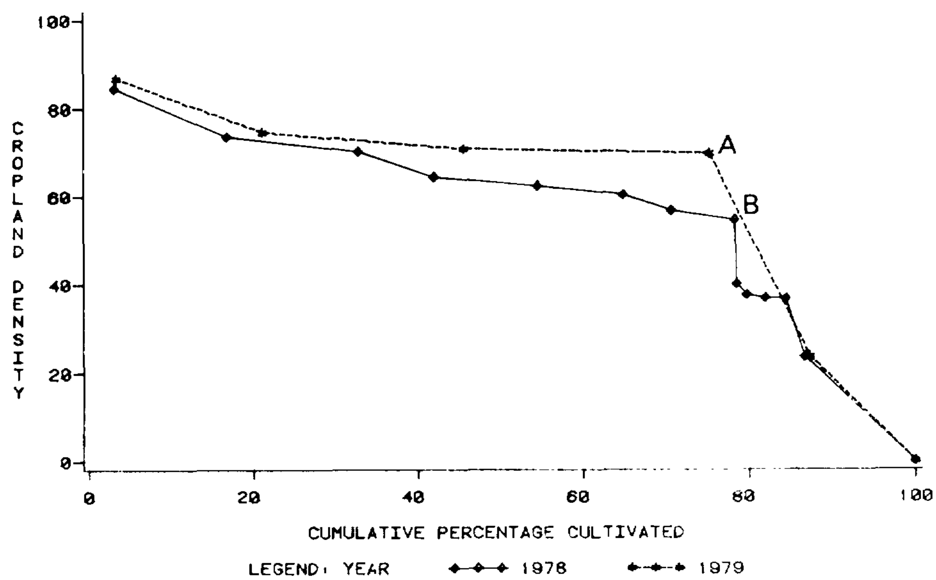
The allocation of land to the various land uses improved dramatically from the old to new frames. This reallocation reflects not only the real changes in land use, but also the increased ability to classify the land use accurately. This increased accuracy is a result of improvements in materials and techniques which are now used in stratification. For example, stratifiers were better able to distinguish land use patterns by using current LANDSAT imagery in frame construction rather than older PI's.

A major use of the SRS area frames is estimating crop acreages. Stratification by land use is used to increase the efficiency of these estimates. Cropland density was used to compare the effectiveness of the stratification for the old and new frames. Figure 1 illustrates the comparison between cropland density and the stratification effectiveness as measured by the cumulative percentage of cultivated land. The cropland density scale for a stratum is a percentage determined by dividing the total cultivated land in the stratum by total land in the stratum.

Figure 1 indicates the cumulative percentage cultivated for a given cropland density. The percentage of total cultivated land that has been included in a strata with at least the given cropland density is the cumulative percentage cultivated. For example, point A indicates that the new frame contains almost 80 percent of all cropland classified into land use strata with cropland density greater than 70 percent.

Point B indicates that the old frame had 80 percent of the cropland stratified into land use strata, some of which did not exceed 60-percent cropland in density. Thus, the plunging of each line represents the loss in stratification effectiveness due to the creation of strata with mixtures of cropland and noncropland. The higher level of the new frame line signifies the increase of cropland included in the desired strata.

FIGURE 1--COMPARISON OF CROPLAND DENSITIES FOR THE OLD AND NEW FRAMES



With effective cropland stratification, segment size changes can provide additional efficiency. For example, the target segment size for all cultivated land in the old frame was 1 square mile (640 acres), while the target segment size for the cultivated land in the new frame was one-half square mile in the orchard stratum and 1 square mile in all other cropland strata. By varying the target segment size and sample size in the cultivated land strata, enumerated land area in the new frame was one-third less than in the old frame. Similar reductions were obtained in the noncropland strata as well. Overall, the enumerator's workload decreased substantially. Just over 10,200 total tracts were enumerated in 1978, but only 6,900 were enumerated in 1979. Still, gains in precision due to stratification will more than offset the losses due to a smaller sample size.

Since the area frame is intended to be a multipurpose frame, SRS is interested in the estimate related to specific agricultural land uses rather than in the estimates of general land use. Table 2 shows the comparison between the AFAP, JES estimates and errors, and the official estimates of the USDA's Crop Reporting Board for selected crop items for the old and new California area frames. 12/ The estimates from the combined AFAP and large farm estimates and the estimates from the 1979 JES would be identical had we not made any corrections to reduce nonsampling errors or introduced a small bias due to truncation.

The following analyses are based on comparisons of the old and new frames shown in table 2 under the headings "June Enumerative Survey, 1978" and "Area Frame Analysis Package," respectively. Pasture and nonagricultural land estimators are generally more precise in the new frame. They are also more precise than the crop estimators since they are not as rare. The coefficients of variation (CV) for the selected crop items ranged from 7.1 to 27.9 percent in 1978 and from 7.6 to 41.7 percent in 1979. For estimates of major crops to be considered usable, the CV's should be in the 3- to 12-percent range. Only cotton met the criterion for accuracy of major crop estimators, although there was a smaller range of CV's for the old frame estimates. Not only cotton but also rice and wheat met this criterion in the new frame. Even though the CV's for rice and

12/ The JES estimates are only one source of indications used by the board in setting the official estimates. Other sources included special surveys which were tailored for specific crop estimates. These special surveys provided estimates for rare items with greater accuracy than possible using a multipurpose survey such as the JES.

Table 2--Comparison of estimates and errors for selected land use and crop items for Califonia, 1978-79 1/

Item code and description	June Enumerative Survey				Area frame analysis package		Crop Reporting Board			
	1978		1979				1978		1979	
	Area frame expansion	Coefficient of variation	Area frame expansion	Coefficient of variation	Area frame expansion	Coefficient of variation	June 1	Final	June 1	Final
	1,000 Acres	Percent	1,000 Acres	Percent	1,000 Acres	Percent	---1,000 Acres ---			
840 All land	91,173	7.4	97,736	.97	97,453	.63	NA	NA	NA	NA
842 Pasture	15,969	13.7	14,776	8.0	14,675	8.0	NA	NA	NA	NA
846 Nonagri. land	31,409	23.7	47,036	3.2	47,153	3.1	NA	NA	NA	NA
511 Wine grapes	NA	NA	364	17.3	361	17.3	318	314	311	302
512 Table grapes	NA	NA	56	28.5	55	28.5	62	62	61	62
513 Raisin grapes	NA	NA	253	15.0	250	15.0	241	240	241	240
514 Navel oranges	NA	NA	165	33.0	164	33.0	111	111	110	111
515 Valencia oranges	NA	NA	83	20.5	82	20.6	72	75	73	74
516 Almonds	NA	NA	347	14.0	345	14.0	300	304	321	323
517 Walnuts	NA	NA	284	19.0	282	18.5	180	180	182	180
524 Cotton	1,430	7.1	1,649	7.6	1,641	7.6	1,420	1,480	1,680	1,650
535 Barley	1,195	13.4	778	13.5	767	13.5	1,100	1,100	900	900
540 Winter wheat	582	14.2	740	10.4	734	10.4	650	650	820	820
552 Potatoes	53	27.9	59	32.1	58	32.3	NA	NA	NA	NA
553 Durum wheat	81	19.2	65	37.8	64	37.8	120	120	47	47
570 Sorghum 2/	116	25.0	91	41.1	91	41.7	210	210	180	180
574 Tomatoes	242	14.6	258	16.7	257	16.7	235	235	255	255
605 Rice	461	13.3	448	11.6	446	11.6	460	493	535	525
691 Sugar beets	202	14.4	284	16.7	283	16.7	215	203	215	225

NA = Expansions or estimates not available.
1/ Estimates include large-farm data which is completely enumerated.
2/ Processing tomatoes only, fresh market acreage not included.

wheat were lower in 1979, a comparison with estimates prior to 1978 would produce different results if the acreage of a crop changed noticeably over time. Comparisons were not available for grapes, oranges, or tree nuts which were not estimated in 1978. Although sampling errors were not substantially reduced for the major crops, the new frame is considered more efficient because there was a sizable reduction in sample size and workload.

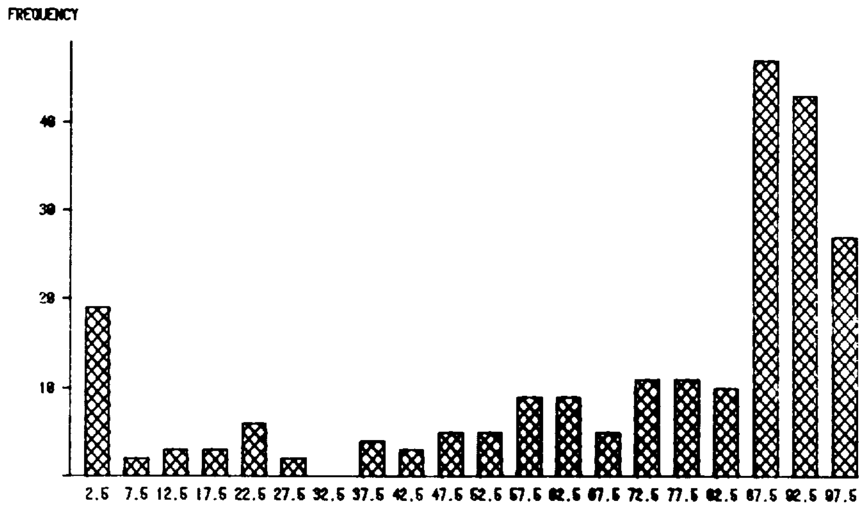
In summary, the CV's for some of the major crops (cotton, rice, and wheat) in the new frame remained unchanged or dipped slightly from the old frame. The CV's for the remaining crops increased but posed no serious problems. However, by reducing sample size and including smaller segments in the orchard stratum, the overall workload dropped by nearly one-third. The inclusion of the orchard stratum now permits estimates for additional commodities of interest (grapes, oranges, and tree nuts), but their CV's are not yet within usable limits. Nevertheless, additional reduction in the CV's are possible through the use of alternative sample allocations which will be discussed in the "Design Evaluation" section.

STRATA ANALYSIS

The percentage of land cultivated is the usual stratification variable used in an SRS area frame. In California, a more efficient sample estimate for important crops was obtained by constructing three crop-specific strata. Crop-specific strata (17 or 19) included land which was 50 percent or more cultivated if 10 percent or more of the land under intensive cultivation was used for either vegetables or fruit crops. Otherwise, this land was put into the general crops strata (13). To see how well the sampled segments conformed to strata definitions, the percentage of cultivated land for each segment in the cropland strata was computed. Frequency distributions of the percentages of cultivated land were prepared for each stratum. Figure 2 shows this distribution for stratum 13, the general crops stratum. Nearly 80 percent of the segments conform to the stratum definition of 50 percent or more cultivated land. The frequency distributions for strata 17, 19, and 20 are included in the appendixes.

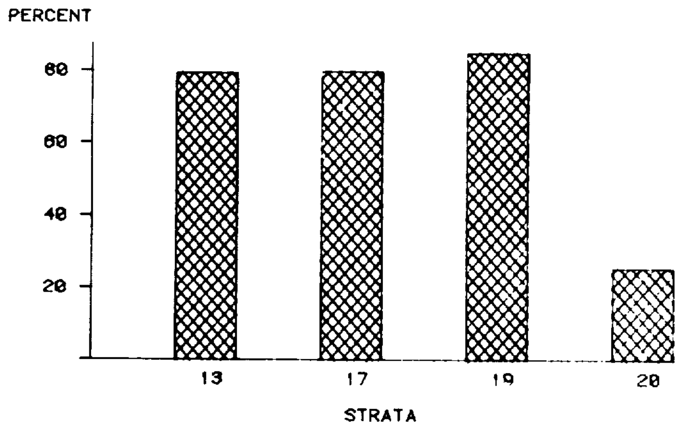
Figure 3 shows the percentage of segments conforming to strata definitions for the cultivated land strata. Segments conformed quite well to the strata definition in the intensively cultivated strata (13, 17, and 19). At least 80 percent of the segments met the definitions. However, only 25 percent of the segments met the stratum definition of 15 to 50 percent cultivated in stratum 20 (extensive cultivation).

FIGURE 2--NUMBER OF SEGMENTS BY PERCENT CULTIVATED LAND
FOR STRATUM 13



NOTE: EXCLUDES SEGMENTS CONTAINING LARGE FARM TRACTS

FIGURE 3--PERCENTAGE OF SEGMENTS CONFORMING TO STRATA DEFINITIONS
FOR CULTIVATED LAND STRATA



NOTE: EXCLUDES SEGMENTS CONTAINING LARGE FARM TRACTS

This lack of conformity to the stratum 20 definition should have been anticipated since this stratum was basically a catchall stratum (land areas that did not meet the definitions for other strata were put into stratum 20 by default). The cultivated land in this stratum tends to be clustered near intensively cultivated land in other strata. This clustering made the formation of homogenous segments virtually impossible in many of the count units in stratum 20. Even though the land, when taken as a whole, fits the stratum definition, the selected sample units often contained either all cropland or no cropland, posing a serious problem which will be addressed further in the Error Review Section.

Stratum 31 and stratum 41, the agri-urban and range strata, respectively, had few segments with significant amounts of cultivated land. As expected, virtually no cultivated land existed in strata 32, 43, 44, 45, or 50. Unfortunately, the few segments containing cultivated land in the range, urban, or nonagricultural strata contributed most heavily to the sampling error.

Table 3 shows the distribution of the percentage of estimated total acres by stratum for selected land uses. Slightly over 40 percent of the fruit crop estimate came from the fruit crop stratum (17), and nearly 25 percent of the vegetable estimate came from the vegetable stratum (19). These results indicated that members of the stratification group were able to identify and place specific crops into their proper stratum.

Table 4 displays each stratum's percentage of the total estimated acres for specified crops. The city, desert range, grazing, and nonagricultural strata made no contribution to these estimates, therefore, these strata were excluded from the table.

Table 5 shows the percentage of the crop estimate and the percentage of the variance of that estimate for each stratum. In general, the specialized strata contributions to the estimates emerged as expected, especially for crops easily identified in the aerial photography. For example, 92 percent of the estimate and 88 percent of the variance for rice came from stratum 13. The stratification was also effective for estimates of table and raisin grapes as well as tomatoes, although the geographic clustering of the crop and the use of paper strata added to the efficiency of stratification. Notable exceptions included navel oranges, cotton, and barley. Fifty-five percent of the navel orange estimate came from the fruit stratum, and over three-quarters of the variance for navel oranges came from the agri-urban stratum 31. Citrus groves often are scattered throughout urban areas and

Table 3--Percentage of estimated acreage for selected land use groupings, by stratum 1/

Strata	Land use groupings			
	Vegetable crops	Fruit crops	General crops 2/	Cultivated land
	Percent			
Large farms	2.9	6.4	77.5	86.8
13	4.6	6.5	58.4	69.5
17	5.1	40.9	24.7	70.7
19	24.9	3.3	46.4	74.6
20	1.8	5.0	16.8	23.6
31	0.3	4.9	1.9	7.1
41	0.1	0.5	2.4	3.0
43	0.	0.	.1	.1
44	0.	0.	1.2	1.2
Statewide 3/	1.09	2.29	6.55	9.93

1/ Excludes strata 32, 45, and 50, which made no contribution to estimated totals.
2/ Computed as the difference between cultivated land and vegetable and fruit crops.
3/ Percentage of total acres for the State.

consequently are difficult to place in the fruit stratum. In such cases, the breakdown of count units is important. Every effort must be made to make the count units homogenous so that one or two segments will not account for a large part of the variance in a given stratum.

Crops thinly dispersed throughout a wide area caused problems in the range stratum 41. For example, only 7 percent of the cotton estimate and 29 percent of the barley estimate came from stratum 41, but 42 percent of the variance of the former and 52 percent of the latter came from that stratum. The sorghum and winter wheat percentages for the estimate were 34 and 9 percent in stratum 41, while the percentages for the variance were 69 and 24 percent.

In summary, the stratification group followed the strata content guidelines and boundary selection rules quite well. The analysis pointed out the potential for increased precision through the use of crop-specific strata, altered boundary selection rules, and changes in count unit breakdown.

Table 4--Percentage of total acres for selected crops, by stratum

Crop	Strata						
	13	17	19	20	31	41	Large farms
	Percent						
Wine grapes	1.1	6.1	1.8	0.8	0.1	0.1	1.2
Table grapes	1/	1.4	-	1/	-	-	1.5
Raisin grapes	.1	6.5	.3	1/	1.0	-	1/
Navel oranges	1/	2.7	1/	.4	2.0	1/	.1
Valencia oranges	1/	1.9	-	.2	-	1/	-
Almonds	1.7	7.0	.4	.3	.6	-	1.6
Walnuts	1.6	4.0	.4	.9	.7	1/	-
Cotton	16.0	9.3	12.5	2.3	.7	.4	41.3
Barley	3.6	1.8	6.5	3.4	.1	.8	8.3
Winter wheat	6.1	2.6	6.8	2.6	.3	.2	10.4
Potatoes	.3	1.1	.1	0.1	-	1/	.4
Durum wheat	.6	-	.6	0.5	-	-	-
Sorghum	.7	.5	-	0.3	-	.1	-
Tomatoes	1.9	.7	6.6	-	-	-	1.1
Rice	9.9	.3	1.1	-	-	-	.6
Sugar beets	3.0	1.1	4.2	0.4	-	1/	1.2

- = zero crop estimate.
1/ = less than 0.1 percent.

ERROR REVIEW

One of the functions of the area frame analysis package is to identify sources of survey error. Once these sources have been identified, the statistician can recommend changes which would reduce survey error. These recommendations could include tightening quality control procedures, allocation changes, variation in the survey design, or altered construction procedures.

AFAP is used to identify and classify sources of survey error in the California frame. When the AFAP master file was created from the JES edited data tape, each observation passed through edit checks. Data for any segment which failed an edit check was placed in an error file along with an identifier describing the reason for the failure. For example, the data for a segment was placed in the error file when observed acreage was not within 0.90 to 1.10 of the planimetered acreage. A portion of the error file for California is included as an example in the appendix.

Table 5--Crop acreage and variance estimates compared as a stratum percentage of the total estimate 1/

Crop and estimates	Stratum					
	13	17	19	20	31	41
	Percent					
Wine grapes:						
Estimate	13	57	12	11	1	6
Variance	11	30	11	41	1	6
Table grapes:						
Estimate	1	97	-	2	-	-
Variance	1	99	-	1	-	-
Raisin grapes:						
Estimate	1	86	3	1	9	-
Variance	1	78	2	1	19	-
Navel oranges:						
Estimate	1	55	1	12	29	2
Variance	1	15	1	7	78	1
Valencia oranges:						
Estimate	1	79	-	14	-	7
Variance	1	82	-	12	-	5
Almonds:						
Estimate	20	69	2	4	4	-
Variance	24	60	3	5	8	-
Walnuts:						
Estimate	24	47	3	6	6	3
Variance	19	25	1	42	11	2
Cotton:						
Estimate	45	21	19	8	1	7
Variance	20	9	11	16	2	42
Barley:						
Estimate	20	8	20	22	1	29
Variance	8	2	11	27	1	52
Winter wheat:						
Estimate	36	13	23	18	1	9
Variance	17	6	19	32	1	24
Potatoes:						
Estimate	22	63	2	9	-	3
Variance	9	82	1	7	-	1
Durum wheat:						
Estimate	36	-	23	41	-	-
Variance	15	-	34	51	-	-
Sorghum:						
Estimate	30	18	-	18	-	34
Variance	8	4	-	19	-	69
Tomatoes:						
Estimate	31	9	60	-	-	-
Variance	34	5	61	-	-	-
Rice:						
Estimate	92	2	6	-	-	-
Variance	88	2	10	-	-	-
Sugar beets:						
Estimate	45	13	35	7	-	1
Variance	24	7	51	17	-	1

- = Zero crop estimate

1/ Percentages across strata may not add to 100 due to rounding.

The error sources were analyzed to determine whether the resulting inaccuracy would lead to a sampling or a nonsampling error. The sources of error are broken down by stratum in table 6 and are defined in the sections on sampling and nonsampling error.

Sources of error were determined by a statistical analysis of the enumerated data and a review of the survey materials for the segments in the error file. These materials included photo indexes, frameworks, photo enlargements, quandrangle maps, and questionnaires.

Sampling Errors

Segment sizes and their distributions were examined to determine indications of increased sampling error due to problems with frame construction. When the segments are formed, an attempt is made to keep all segments approximately equal to a target segment size which is set for each stratum.

Table 6--Number of observations by source of error for selected strata, 1979 California area frame 1/

Source of error	Strata								
	13	17	19	20	31	32	41	44	Total
	Number								
Sampling errors:									
Count unit breakdown	3	3	1	6	2	-	6	-	21
Stratification efficiency	2	3	-	2	-	-	-	-	7
Nonsampling errors:									
Mapping	2	3	4	-	2	1	1	-	13
Not planimetered 2/	11	2	4	6	1	1	5	-	30
Measurement error	5	2	-	-	3	-	5	-	15
Enumeration	2	2	2	1	2	1	3	-	13
Data handling	-	2	1	-	1	-	3	1	8
Total	25	17	12	15	11	3	23	1	107

-= No error observed.

1/ Strata 43, 45, and 50 were excluded because no observations appeared in the error file. Observations that could not be categorized uniquely were placed in categories which could be applied. Thus the total count of observations exceeds the number of observations in the error file.

2/ Due to time constraints some segments were not planimetered prior to the 1979 survey. These observations only represent nonplanimetered segments which failed an edit check.

Since the area frame sample segments are selected with equal probability, the distribution of segment sizes would ideally be unimodal with a mean and mode about equal to the target segment size. Table 7 summarizes the properties of the distribution of segment sizes for the 1979 JES sample.

The mean segment size was significantly different from the target segment size at the 95-percent level for strata 17 and 43, and no significant differences were detected for the remaining segments. About 7 percent of the segments in stratum 13 contained large farms, but for these segments the large farm tracts were excluded from the segment summary. Thus, the effect of these large farms on the true average size for stratum 13 is unknown, and a test of significance cannot be adequately assessed. The distributions of the observed segment sizes for the various strata tended to be non-normal and skewed. Strata having the smaller target segment sizes tended to be positively skewed while the larger range segments were negatively skewed. The deviation of the sample mean from the target segment size appeared to be mainly a result of current methodology for constructing frames in which emphasis is placed on using "good" boundaries.

Table 7--Properties of the distribution of segment sizes, by stratum 1/

Stratum	Number of segments 2/	Mean segment size 2/	Target segment size	Skewness 4/
13	224	634.9	640	positive
17	238	3/ 326.2	320	positive
19	79	657.1	640	positive
20	118	641.3	640	positive
31	40	161.7	160	-
32	10	69.6	64	positive
41	100	2,536.1	2,560	-
43	20	3/ 2,483.4	2,560	negative
45	8	2,512.1	2,560	negative
50	8	1,259.1	1,280	-

- = measure not statistically significant.

1/ Excludes stratum 44 because sample selected with probability proportional to size.

2/ Segments containing large farms excluded from computations of means.

3/ Statistically different from target segment size at the alpha = 0.05 level.

4/ Measures are statistically significant at the alpha = 0.10 level.

Another source leading to segment size deviations was the rounding error which occurred when a count unit was divided into a distinct number of segments. For example, a count unit of 5.4 square miles was assigned five segments when the target size was 1 square mile. Each of the segments would ideally be approximately 1.08 square miles, or 51 acres larger than the target size. If the segment size variance was not equal to zero then deviations in segment size would contribute to sampling error.

The attempts to keep segment sizes equal appear to work fairly well. The restriction of "good" boundaries might be loosened in nonagricultural areas to improve the distribution of segment sizes and help reduce the sampling errors for individual crop estimates. However, before significantly altering restrictions on boundaries, additional research of the effects on the livestock estimates should be undertaken.

Sampling error was also increased by the improper breakdown of count units. This improper breakdown occurred when too much emphasis was placed on having "good" segment boundaries. Problems in the count unit breakdown also occurred when a count unit was split. These factors contributed to deviations in segment size and problems with homogeneity of crop content among segments within the count unit.

Stratum 20 and stratum 41 contained over half of the observations for which the count unit breakdown was a source of survey error. In these strata, "good" segment boundaries were hard to find because there was a scarcity of roads, major rivers, or other visible landmarks to use as the boundaries. Conversely, alternative boundaries were available in strata 13 and 17, areas of intense cultivation. However, the emphasis on "good" boundaries again resulted in deviations from the target segment size.

Although the deviations in segment size can lead to sampling error, the homogeneity of crop content appears to be the most critical element in count unit breakdown. For the count unit breakdown to be most efficient, one must ensure that the difference between the acreages of individual crops or crop types from segment to segment in a count unit are as small as possible. For example, if a count unit had 500 acres of rice and was to be broken down into 10 segments then each segment within the count unit should have contained approximately 50 acres of rice. Other crops should be apportioned in a like manner. Although, in reality, it would be virtually impossible to put equal amounts of each crop in the segments, the difference between crop content of segments within a count unit should be minimized as much as possible.

Another source of sampling error was inefficiencies in stratification. This type of error resulted from land area being allocated to a particular stratum when an alternative and better allocation could have been made. Inefficiencies in stratification cannot be identified directly by the analysis package, but areas of possible misstratification were discovered in the process of reviewing the photo indexes for the segments in the error file. These stratification problems were most acute where intensive and extensive cropland met. For example, in California, areas of dense cropland and small pockets of fruit crops were included in strata otherwise devoid of cultivation. Segments which fell in these areas contributed disproportionately to the sampling error. This problem could have been avoided by using the procedure in figure 4.

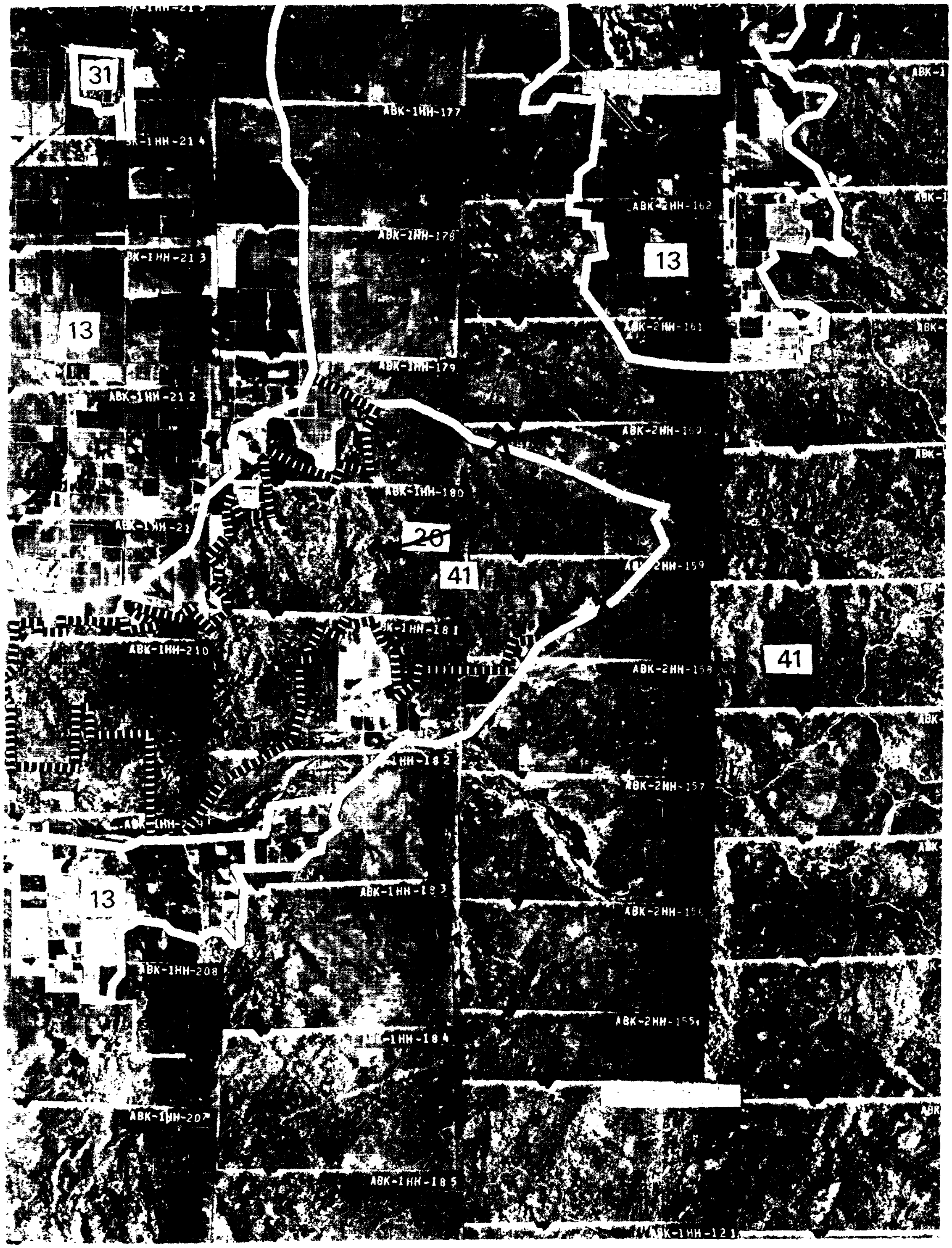
Notice that the canal and roads that were used as strata boundaries do not include all the intensive cropland. The stratifier had to decide whether to use roads as stratum boundaries (solid lines) or to find other boundaries (dashed lines). The former resulted in putting small pockets of dense cropland into an extensive cropland strata while the latter would have included the cultivated land and a small amount of uncultivated land in the intensive cropland stratum. In this example, overall efficiency would best be achieved by "misstratifying" the small amount of noncropland into the intensive cropland stratum rather than "misstratifying" cropland into a low density cropland stratum. Notice how the strata homogeneity was increased by eliminating the highly variable stratum 20 and creating the very homogenous strata 13 and 41.

The outdated photography may have contributed to the stratification error because land use could have changed dramatically in areas near the edge of the strata. We feel that reliance on LANDSAT imagery in constructing the strata and count unit boundaries is necessary to increase the efficiency of stratification.

<u>Nonsampling Errors</u>

Nonsampling errors occurred in mapping, planimetering, and enumerating, three sources sometimes difficult to distinguish from one another. For example, consider an observation that exceeds the edit limits for the comparison of reported acreage to planimetered acreage. A determination whether this edit failure was caused by mismeasurement of planimetered acreage, an error in mapping, or an error in enumerating the segment could not be made without reviewing questionnaires or photo enlargements. However, these items are maintained at the State office and are not easily available for review by statisticians in Washington, D.C. In future analyses, a State survey statistician should review segments which exhibit errors so that the cause can be accurately determined.

Figure 4--Example of procedure to increase efficiency of strata.



Mapping errors could contribute to nonsampling inaccuracies when mistakes are made transferring segment boundaries from the PI's to the photo enlargements (or quadrangle maps). Segment area could be mismeasured by inaccurate planimetering. These errors in planimetering result from ambiguous segment boundaries, distortions in photography, or lack of planimetering experience. Enumeration errors result when segment boundaries are misinterpreted or when tract acres for a nonrespondent are estimated inaccurately. 13/

Planimetered acres may be linked to nonsampling errors since the value is used as a quality control variable. Thus, State survey statisticians may consider this value to be "truth." The total-acre value may be edited to meet the error limitations specified in the JES Supervisory and Editing Manual when discrepancies appear between the total acres reported (sum of the tract acres for a segment) and planimetered acres. Thus, the use of target segment size as an estimate for planimetered acres for some of the segments in California might have contributed to nonsampling error if the estimated value differed from the true acreage of the segment.

Data handling was another source of nonsampling error. These errors included problems in editing, keypunching, and questionnaire handling. Not all data handling errors are discovered and corrected, although JES data are passed through a survey edit system. The JES edit system checks observations primarily at the tract level. Hence, questionnaires may be misplaced and data mispunched with no apparent problem. Eight data handling errors were discovered in the California data set.

Wherever possible, these errors were verified and corrected. The net effect of the error correction was to reduce the coefficient of variation of the estimate for all land from 0.97 to 0.63. Consequently, the data set to be analyzed by AFAP has a reduced potential for bias since nonsampling error has been reduced.

In summary, an error file obtained by editing the JES data file for the 1979 survey in California was used to help identify the sources of error which contributed to the total survey error. The survey error was made up of both sampling and nonsampling inaccuracies. The nonsampling errors were corrected to create

13/ There are other types of enumeration errors such as data incorrectly reported or incorrectly entered on a questionnaire. AFAP is not designed to detect these types of enumeration errors.

a data set with reduced total error. This "clean" data set was then used for the bulk of the analyses.

DESIGN EVALUATION

Two topics will complete the evaluation of the California area frame: the sources of sampling variation that are a result of the design and the methods available to control the variance. The sources of variation include: replication and rotation groups, paper strata, sample unit definition, stratification, and sample size and allocation.

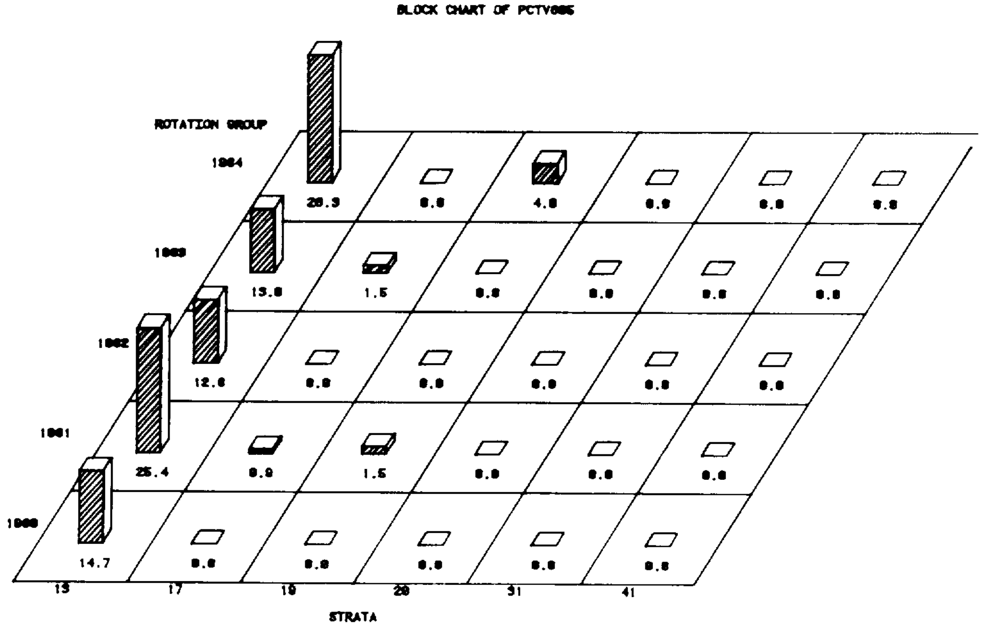
When discussing the impact of the sources of variation, two-way histograms (block charts) are used showing the percentage of the crop estimate by rotation year and stratum. Rice and cotton are the crops used in the discussion (figs. 5 and 6). The block charts for other crops are included in the appendices. Inspection of the between-column block heights shows the effects of stratification, while the between-row heights show the impact of the rotation groups. To some extent, the effect of replication is also shown since the rotation groups are comprised of independent, mutually exclusive groups of replications.

Statistical Design

In figure 5, the block chart for rice, notice the rotation differences in stratum 13 where the various rotation groups contributed 13 to 26 percent of the total estimate. However, few statistical differences were detected between replications in an analysis of variance (ANOVA) because of the large estimated variance within the replications. This large variance is due to effective paper stratification and the use of a simple random sample (srs) estimate of variance within replication. Yet, as with rice, instances exist where the rotation of sample segments can lead to a 10 percent or more change in the level of the estimate (refer to appendix figures 4 through 13). This variation in replication level justifies the need for increased research on the use of ratio estimators with segments which have been in the sample for more than the current year.

The variation among replications results from the breakdown of count units into segments. The use of old photography and a minimal effort to identify individual crops cause most of the variation. We found that the count unit breakdown does not preserve homogeneity of crop content between segments. The result is a nearly binomial approximation of the variance for the individual crops. The segments in a paper stratum that contain a crop each have about the same acreage, or they have none of the crop at all. Additionally, when a major crop is present in a California segment, the crop acreage is often a

FIGURE 5--PERCENTAGE OF ESTIMATED RICE
BY ROTATION GROUP FOR SELECTED STRATA, 1979

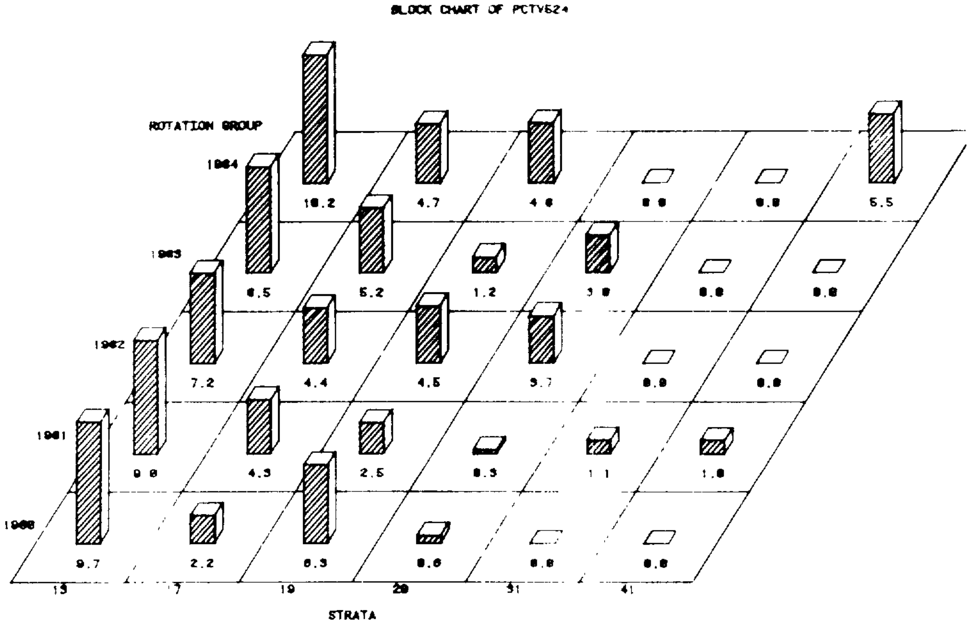


large proportion of the segment land area. Appendix figures 1 through 3 show the extent of this binomial effect. Referring to figure 2, the proportion of cultivated land to total land is one-quarter or less in 15 percent of the segments, while the proportion is three-quarters or more in 62 percent of the segments. This shows the polarity of cultivated land density even in a stratum which is very dense in cultivated land. This polarity becomes more exaggerated for any one of the specific crops which make up cultivated land because a crop is more of a rare item.

Since the segment proportion of each crop is often close to 0 or 1, we are left with a situation very similar to a binomial distribution. Also, the paper strata sample size is usually less than 10. Thus, the estimator is not close to being normally distributed, and small increases in sample size have virtually no variance reduction benefits. 14/

14/ Note that an increase of one sample in a paper strata may require a land use strata sample increase equal to the number of paper strata. That is, a replication, which is often more than 20 segments, must be added if the design remains unchanged.

FIGURE 6--PERCENTAGE OF ESTIMATED COTTON
BY ROTATION GROUP FOR SELECTED STRATA, 1979



Estimation of totals in this type of population results in the standard error of the paper strata estimator being almost equal to the range of the individual sample expansions. 15/ In many of the paper strata there was only one segment with the crop, or one segment had considerably more of the crop than any other in the paper stratum. In such cases, the paper stratum standard error is dominated by this large observation.

The way to reduce both paper strata and replication variances is to concentrate more on breaking down count units. The cause of these problems and some remedies will be discussed further in the next section.

Paper stratification was effective for most crops, especially cotton and rice. In each land use stratum, the crops appeared in fewer than one-half of the paper strata. This concentration of the commodity results in precision above that of simple random sampling. More paper strata appear in California than necessary since most of the crops are contained in five or six adjoining paper strata.

15/ See appendices (twovar module). The column labeled 524RNGE is the range of expansions in the paper strata. PPSE is the paper strata standard error.

An earlier study by Pratt (7) examined sample designs having paper stratification. This work, which was done with the Nebraska area frame, suggested that the best design has about as many paper strata as replications, a square design that appears indigenous to Nebraska. A more general approach features constructing paper strata such that the number of paper strata containing a crop is minimized while the density of the crop in these paper strata is maximized. Fewer paper strata allows for more replications, since the fixed sample size is equal to the number of paper strata times the number of replications. Fewer paper strata also means increased flexibility in sample allocation. Thus, the process of creating effective paper strata is really a weaker version of crop-specific stratification for geographically isolated commodities. Although less effective than crop-specific stratification, paper stratification is consistent with the desire to make the distribution of the estimator approach normality. Also, paper strata often improve the estimate of crops which could not be distinguished by photo interpretation during the stratification process.

Design Implementation

The effectiveness of a statistical design is related to implementing the design given specific data collection requirements. An SRS area frame must be general purpose, that is, the frame may be used to estimate a wide variety of agriculturally related parameters. When the earliest of these general purpose area frames was developed for the West and Midwest, the concept of agricultural intensity was used for stratification and count unit breakdown. Since most of these States have a limited crop variation between fields in an area, the stratification was generally efficient. However, inspection of the aerial photography for California showed that agricultural density is not highly correlated with any particular crop. Thus, the stratified design defined solely by agricultural density is not always the most efficient way to create a general purpose frame because the sample size cannot be altered to reduce the sampling error of a particular item without affecting the variability of the estimates of other items. Numerical examples of the effect of alterations in the sample allocation will be presented after discussing crop-specific stratification.

The availability of LANDSAT imagery made crop-specific stratification possible in California. The use of a limited number of strata for specific crop groups, namely the fruit and vegetable strata, was highly effective in California (see tables 3 and 4 and the block charts in appendix figures 4 through 13). Yet, many of the crops--rice, grapes, and cotton in particular--are geographically limited and are recognizable in the stratification materials. The use of some crop-specific strata for major items like these provides more efficiency as well as versatility. Research on the use of prior survey data to determine the usefulness of crop-specific stratification in a new frame is currently active in the SFDS section.

Finally, the process of breaking down count units needs much attention. Since stratification is a major portion of the frame development cost and the main control of the population variance, any losses in efficiency due to a nonhomogenous breakdown of count units is counterproductive and should be minimized.

We have already explained how the nonhomogenous breakdowns led to the binomial-like variances found in the analysis. Apparently, the emphasis on keeping segment size within tolerable limits has overshadowed the real need to equalize the amounts of each crop in as many, if not all, of the predefined numbers of segments within the count unit. This is why the variances are hard to control. For example, 5 percent of the cotton estimate came from one segment in stratum 41 which had 640 acres of cotton. With a more homogenous count unit breakdown, the selected segment would have contained, at most, 160 acres of cotton. This one change would reduce the CV of the cotton estimate for the State by 17 percent. Similar problems exist for sorghum in stratum 20 and stratum 41. In developing area frames, less emphasis on segment size limits is necessary to attain homogeneity of crop acreages, resulting in a frame with as little variability as possible between segments within a stratum. Thus, smaller sample sizes will efficiently estimate stratum totals.

We went to the framework and simulated breaking down count units on a crop-specific basis to determine the implementation effects on overall variability. Sample segments showing a disproportionate amount of a single crop were identified. If the segment appeared to be very different from other segments in the count unit, the count unit was broken into the desired number of sample units to reflect the proposed procedure. A simulated segment was then used in the summary in lieu of the original segment.

Table 8 displays the amount of reduction in CV obtained with this minimal effort toward a homogenous count unit breakdown for specific crops. In the simulation for each crop, four segments, at most, had acreage changes based on a limited inspection of the frame materials. Even with so few changes in the simulation, the standard error reductions are noticeable, especially for the crops which had larger CV's in the new frame. Using LANDSAT color imagery as a check on the content of other count units would surely provide simulations with further CV reductions, especially when used in the highly variable strata like 20 or 41.

The following changes to the count unit breakdown are recommended:

1. LANDSAT imagery should be used to identify changes in the count unit land use as seen on the PI. Interpret the new land use, if any, and note it on the PI.
2. Groups of fields that have similar cropcover should be identified on the PI. This is especially easy for grapes, rice, and tree crops.
3. These groups of fields should be split into as many different count units as boundaries and reasonable segment size limits permit.
4. A segment that obviously does not come close to meeting the stratum definition of crop content should be avoided.
5. If a block of fields with a particular crop use is so large that segments must be made up entirely of fields in that block, those segments must be made smaller than the target segment size for the stratum. As a rule, do not create a segment for which a particular crop use is more than twice the area which would occur if the cropland could be equally divided into all the segments in the count unit.

Table 8--Simulated efficiency of a homogeneous count unit breakdown for specific crops 1/

Crop	Coefficients of variation		Standard error reduction
	JES sample	Simulated sample	
	Percent		
Cotton	8.9	8.3	7
Sorghum	41.7	29.6	29
Sugar beets	17.0	15.8	7
Walnuts	18.7	15.5	17

1/ Excludes large farm estimates.

Sample Size and Allocation

The JES is designed to provide a relative standard error of 2 to 4 percent for major items on a national level and 3 to 12 percent on a State level. 16/ The sample size and allocation for a State generally depend on the desired precision and the available budget (cost). Basically, the number of tracts is a measure of the survey cost. Therefore, the statistics of interest for sample size and allocation would be the number of segments per stratum, the average number of tracts per segment by stratum, the cost difference per tract by stratum, and stratum variances.

The Methods Staff of SRS uses Neyman allocation for a simple random sample to obtain approximations to the estimators of sample size and stratum allocation. To apply the allocation algorithm, stratum variances are estimated using the variance formula for simple random sampling. 17/ Unfortunately, the robustness of these approximations was questionable; therefore, new estimators must be developed.

To develop the exact estimator of the optimal allocation, consider an approach which makes no assumptions about approximations to the design variances. The only assumption necessary concerns the costs. Since no time statistics are kept by tract, only an approximation to the cost per tract can be made. The frequency of tracts having livestock was used to calculate a relative cost statistic. 18/ The rationale was that livestock presence would require additional time to obtain the required information. Let the cost of completing a tract questionnaire with livestock be $C = 1$ cost unit, while the livestock portion of the cost is $0.2C=0.2$. Now let the proportion of tracts without livestock be L_h in stratum h. Hence, the cost of a tract in stratum h is

$$C_h = (1 - L_h)C + L_h(C - 0.2C) = C(1 - 0.2L_h).$$

Also, we stated that $C = 1$, therefore

$$C_h = 1 - 0.2L_h.$$

16/ For some items, mainly livestock items, a list frame is used in conjunction with the area frame (multiple frame surveys) to achieve the desired CV's.

17/ The allocation algorithm used a convex programing approach to Neyman allocation for multiple items (6).

18/ More detailed relative costs can be developed by including hogs and cattle separately while accounting for the frequency of resident farm operators. The point is that differences between strata costs will affect optimal allocation. A more detailed cost analysis is planned in the New York Area Frame which is tentatively scheduled for publication in 1981.

Let P_h be the number of paper strata and T_h be the average number of tracts per segment in stratum h. Then the cost per replication in stratum h is $D_h = P_h \times C_h \times T_h$. The S_h for the commodity listed is the standard deviation between replications in the sample.

The problem is to minimize the variance V subject to the 1979 fixed cost C` where

$$V = \Sigma\, n_h^{-1}\, S_h^2 ,$$

$$C' = \Sigma\, D_h\, R_h = \Sigma\, D_h\, n_h = 5529,$$

and R_h is the number of replications in stratum h in 1979.

The cost function is the enumeration cost per replication times the number of replications. No travel cost is considered because this is fairly uniform given our density of sampling (1 square mile in an approximately 5-x 8-mile area for intensive agricultural strata) and the daily workload of an enumerator.

Using Lagrange multipliers, we need to minimize

$$\Sigma\, n_h^{-1}\, S_h^2 + \lambda \Sigma\, n_h\, D_h ,$$

where the summation is over the strata being allocated. Differentiate with respect to n_h, sum over the strata, form the appropriate ratios (2), and use the cost function to form the following relationships for stratified estimates of a total using replicated sampling:

$$\frac{n_h}{n} = \frac{S_h\, D_h^{.5}}{\Sigma(S_h\, D_h^{.5})} \quad \text{and} \quad n = \frac{C'\; \Sigma(S_h\, D_h^{.5})}{\Sigma(S_h\, D_h^{.5})} .$$

Thus, we now have the estimator for the optimal allocation according to our design. Cotton is used to show the calculations for this allocation (table 9). Other crop-specific allocations are contained in table 10.

Consider the calculations for the exact estimator of the optimal allocation for cotton. Considerable changes exist in the strata sample sizes (compare the second and last columns), especially stratum 41 where sample size is almost doubled. The overall sample would increase by 18 segments, from 820 to 838 but the number of tracts would increase only 12 from 6,708 to 6,720. The overall "cost" of the survey would increase slightly from 5,529 to 5,558 due to rounding the n_h. Even with this crop-specific approach to the allocation, the standard error of the State level CV drops only from 8.9 to 8.0. That is, the CV of 8.0 is the best we can achieve for cotton using the current procedures and budget.

The above example on reallocation shows how the binomial nature of the data and the small number of replications restrains the ability to reduce the variance noticeably by making small increases in the number of replications. Yet for the 10-percent reduction in CV, we would prefer the use of the proposed optimal allocation to current procedures. When the optimal CV is computed using the data from the previously mentioned simulation of the count unit breakdown for cotton, the CV is reduced from 8.3 to 6.0 percent, a 38-percent CV reduction and another indication of the value of this allocation.

Table 9 helps explain the surprisingly sharp decline in the number of tracts from the 1978 sample to the 1979 sample. Although part of the decline was attributable to a reduced number of segments, the most important factor by far was segment size and sample size by stratum. Stratum 17 segments contain a higher density of tracts than other densely agricultural strata. Thus, the smaller segment size reduced the number of tracts by about 2,381 (240 x 9.92), as was expected. The tract reduction in a fruit-dominated stratum had little effect on the major crop variances. Still, the largest portion of the survey "cost" (2,381 tracts) was spent in this stratum, while fruit crop CV's remained outside usable limits. The cost benefit of such heavy sampling in stratum 17 is therefore questionable.

The benefit of developing the optimal allocation estimator which fits the design may best be seen by examining the reallocation of the sample done by the Methods Staff after the 1979 JES. The survey data was used to compute estimates of the variance by applying the srs formulas. These variance estimates were then used in the convex programing approach to estimate the optimal allocation. The result is an estimated optimal allocation which would slightly decrease stratum 13, decrease stratum 17 by about half, increase stratum 19 by a quarter, and double stratum 41, while the remaining strata keep the same sample sizes (table 10).

The increased allocation in stratum 41 could be expected since the variance of the cattle estimate was used, and the cotton estimate was very imprecise in stratum 41 because of the one

Table 9--JES sample size and cost statistics for sample reallocation 1/

Strata	Sample size	Tracts	Average tracts per segment T_h	Expected segment size 2/	Proportion of tracts without livestock L_h	Paper strata P_h	Replications, 1979 R_h	Expansion factor, 1979	Cost per tract $1-2.L_h = C_h$	Cost per replication $P_n x C_h x T_h = D_h$	$\sqrt{D_h}$	Variance for cotton S_h	$\frac{S_h}{\sqrt{D_h}}$	$S_h \sqrt{D_h}$	$\frac{n_h}{n}$	n_h	Optimal sample size for cotton
				Miles	Percent	Number											
13	240	1,667	6.95	1.00	0.855	24	10	29	0.829	138	11.8	146,391	12,406	1,727,414	0.1239	8	192
17	240	2,381	9.92	.50	.916	24	10	43	.817	194	13.9	126,146	9,075	1,753,429	.0906	6	144
19	80	618	7.73	1.00	.911	10	8	45	.818	63	8.0	162,229	20,279	1,297,832	.2025	13	130
20	120	840	7.00	1.00	.823	12	10	64	.835	70	8.4	196,351	23,375	1,649,348	.2334	14	168
31	40	323	8.08	.25	.920	8	5	369	.816	53	7.2	37,164	5,162	267,581	.0515	3	24
32	10	72	7.30	.10	1.000	2	5	2315	.800	12	3.4	0	NA	NA	NA	NA	NA
41	100	879	8.79	4.00	.799	10	10	104	.840	74	8.6	256,844	29,866	2,208,858	.2982	18	180
43	20	48	2.40	4.00	.979	5	4	199	.804	10	3.1	0	NA	NA	NA	NA	NA
44	25	36	1.44	NA	.583	5	5	2/	.884	6	2.5	0	NA	NA	NA	NA	NA
45	8	12	1.50	4.00	.917	2	4	501	.817	2	1.5	0	NA	NA	NA	NA	NA
50	8	8	1.00	2.00	1.000	2	4	886	.800	2	1.3	0	NA	NA	NA	NA	NA
Total sample	891	6885	7.73	NA	.874	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Total 3/	820	6708	8.18	NA	NA	NA	NA	NA	NA	NA	NA	NA	100,163	8,904,462	NA	NA	NA

NA = Not applicable.
1/ Strata with a zero estimate for the variance are not reallocated.
2/ Sample selected with probability proportional to size - expansion factor is variable.
3/ Total for the reallocated strata (13, 17, 19, 20, 31, 41).

Table 10--Optimal allocations for specific commodities

Strata	1979 allocation	1980 methods staff reallocation	Commodity						
			Cotton	Corrected cotton	Rice	Tomatoes	Raisin grapes	Cattle	Proposed reallocation
			Segment						
13	240	229	192	264	480	360	48	312	288
17	240	130	144	192	48	96	432	72	192
19	80	105	130	180	90	150	50	230	120
20	120	118	168	120	-	-	12	96	120
31	40	-	24	40	-	-	112	144	60
32	10	-	-	-	-	-	-	-	8
41	100	218	180	60	-	-	-	30	60
43	20	-	-	-	-	-	-	-	15
45	8	-	-	-	-	-	-	2	6
50	8	-	-	-	-	-	-	-	6
Total for reallocation 1/	NA	NA	820	820	560	560	720	828	866
Total reallocated 1/	NA	NA	838	856	618	606	654	884	865
1979 coefficient of variation 2/	NA	NA	8.9	8.3	13.4	12.2	10.4	NA	NA
Optimal coefficient of variation	NA	NA	8.0	6.0	10.4	10.6	10.4	NA	NA

\- = optimal allocation not completed.
NA = not applicable or not available.
1/ Segment total for strata being reallocated.
2/ Standard errors computed from replications.

segment discussed earlier. Since the area frame is a form of cluster sampling, items such as cattle are more efficiently estimated by multiple frame estimates. The differences in allocations indicate further need for research when making use of cattle estimates in determining the area frame allocation. In fact, an earlier study (1) did not use livestock variables directly to determine optimum strata boundaries and optimum number of strata. Instead crop variables were used showing the effect of the optimum value on livestock.

Although the optimal allocation is not yet extended to the multivariate situation, the individual commodity allocations provide the basis for a general purpose allocation (table 10). Unless the emphasis is on fruit or nut estimates, the current sample size in stratum 13 is insufficient. The estimators for fruits and nuts cannot be brought to usable precision levels (CV less than 12 percent) without further concentrating the survey resources in stratum 17 and severely reducing the efficiency of the major crop estimates. Hence, a sharp reduction in the sample size in stratum 17 is easily justified. The last column shows a proposed reallocation which would improve efficiencies of the estimators for major crops at the expense of the currently high CV estimators for minor crops. The number of segments would increase slightly, but the number of tracts would be reduced by 1 percent.

This reallocation shows the need to re-evaluate the ideas of strata definitions. Although the bulk of the rice and cotton fall in stratum 13, there is no allocation which can minimize the variance of one crop without increasing the variance of another. This problem may have been avoided by dividing stratum 13 into a limited number of crop-specific strata rather than the large number of paper strata. Stratum 13 had 6 paper strata which contained almost all of the rice, while a different 11 paper strata contained all of the cotton.

For example, consider the following restratification and reallocation of strata 13, 17, and 19 (table 11). The new strata are formed by combining groups of paper strata within a stratum. The original three strata contained 560 segments, but under our suggested allocation, contain only 522 segments. By reallocating segments into the crop-specific strata for rice and cotton, we obtained CV's for both rice and cotton which were at least as low as the CV's for the individual optimal allocation for the crops under the original stratification. Yet, under the original stratification, both low CV's could not be attained from any one allocation. Thus, the crop-specific strata make joint CV reductions attainable--an impossible calculation using a fixed sample size and the original strata. Further, the increase in efficiency was achieved with a

Table 11--Restratified allocation for strata 13, 17, and 19

Crop-specific strata and crops	Paper strata reallocation	Number of paper strata	Number of replications	Total segments
13		Number		
Rice	4-9	5	20	100
Cotton	14-24	11	11	121
Other	1-3, 10-13	7	3	21
17				
Rice/cotton	8,12	2	8	16
Cotton	13-24	12	8	96
Other	1-7, 9-11	10	6	60
19				
Rice/cotton	2, 3, 6	3	9	27
Cotton	7-10	4	18	72
Other	1, 4, 5	3	3	9

7-percent reduction in sample size. Therefore, the CV's can be reduced below 5 percent for cotton and rice, as well as other distinguishable crops such as tree fruits or grapes, if these crops were included as stratification variables. Some crops cannot be identified with current materials available for stratification, but any possible crop-specific stratification of crops with large acreage will give the frame more potential for precision for all crops.

The deviations among the original allocations and the various reallocations indicate the need for a more rigorous approach to setting the design and allocation for new area frames. Strata definitions are set by inspection of frame materials, prior segment data, and county estimates. The allocation to these strata is based on variance approximations which are computed from old frame segments fitting the new strata definitions. To improve the new frame's performance, the old frame segments should be classified into the new frame strata based on their location on the new framework rather than on segment content. The segments can then be arranged based on the county ordering which is used to create paper strata. Various combinations of paper strata and replications can then be simulated. Inspection of the paper strata content can show the need for some crop-specific stratification as was shown for rice. Restratification may be done by inspecting the existing framework and altering the strata code of the desired count units.

This restratified frame can have paper strata and replication combinations simulated to determine the optimal design. Then the variances computed from the optional design may be used to compute the optimal allocation.

Research has indicated that the individual commodity estimates would be more precise with a design containing a reasonable number of crop-specific land use strata, about as many paper strata as replications within each of the crop-specific land use strata, and more replications than there are paper strata in multicrop land use strata.

RECOMMENDATIONS

Analyses show that the new frame for California is more efficient than the old frame. The replicated design allows for easy reallocation of the sample whenever required in order to provide reduced sampling errors or to alter survey costs. Reliable estimates for crops previously having CV's outside the usable range can be achieved through reallocation and changes in the frame construction procedures. The use of specialty strata improves the efficiency of the stratified design.

Some problems turned up in stratification, segment size distributions, and segment content which increased sampling errors. Errors in mapping, enumeration, and summarization led to minor nonsampling errors; however, improved procedures for quality control should lessen their impact on future survey results.

The results of this analysis provided both detailed information on how well the new frame did and pointed out areas where improvements in procedures for constructing area frames can be made. Our recommendations for improving area frame methodology fall into three general categories: frame construction, quality control, and research.

Frame Construction

(1) Compile area frame instructions into a single comprehensive set. When this study was initiated, materials were sketchy, out of date, and scattered among working units. These instructions should be detailed and include not only traditional procedures but also new procedures such as using LANDSAT imagery in constructing a frame.

(2) Stress the concept of segment homogeneity when breaking down count units into segments. Segments in a count unit should have an even distribution of cropland acreage in general and specific crops where possible. LANDSAT imagery

should be used to help determine the distribution of crops. This is especially true in strata 20 through 50 and when the PI's are old.

Quality Control

(1) Continue the development of a quality control (QC) system for all stages of frame construction. Some QC procedures, such as that used in the selection of random numbers, have recently been implemented. However, additional QC measures are needed to ensure procedural consistency.

(2) Digitize segments to obtain consistent values to be used for quality control in the field. Presently planimetered acreage is used. This value has exhibited a high potential for error and can lead to increases in nonsampling errors.

(3) Develop methodology to do post survey QC checks. LANDSAT imagery or 35mm slide coverage from the Agricultural Stabilization and Conservation Service provide two possible tools for the QC checks.

Research

(1) Develop a pilot study incorporating the use of auxiliary information assigned to each count unit as a procedure to improve stratification, increase the life of frames, and reduce frame construction costs.

(2) Explore the development of a system similar to the one used by the Remote Sensing Branch that would produce grey scales (3) for use in breaking up count units. 19/ Also, current research in the AgRISTARS program, which is exploring full frame sampling may produce results useful to count unit breakdown and stratification.

(3) Consider the possibility of plotting count unit boundaries which contain sampled segments on acetate and registering these boundaries on current remotely sensed imagery. These registered plots can then be used to improve the count unit breakup.

(4) Study the possibility of using results of previous surveys to develop the strata for new frames. This study should include methodology for determining the optimal mix of paper strata and replications.

(5) Continue research on optimal segment size and land use characteristics of segments used in the SRS area frames.

19/ A grey scale is each picture element represented by a one-print character whose lightness/darkness corresponds to the reflected energy in a single band as measured by the satellite scanner.

REFERENCES

(1) Ciancio, Nicholas J., Dwight A. Rockwell, and Robert D. Tortora. An Empirical Study of Area Frame Stratification. U.S. Dept. Agr., Stat. Rept. Serv., July 1977.

(2) Cochran, William G. Sampling Techniques. 2nd ed. New York: John Wiley and Sons, Inc., 1963.

(3) Craig M., R. Sigman, and M. Cardenas. Area Estimates by LANDSAT: Kansas 1976 Winter Wheat. U.S. Dept. Agr., Econ. Stat. Coop. Serv., Aug. 1978.

(4) Hanuschak, George A., and Kathleen M. Morrissey. Pilot Study of the Potential Contributions of LANDSAT Data in the Construction of Area Sampling Frames. U.S. Dept. Agr., Stat. Rept. Serv., Oct. 1977.

(5) Houseman, Earl E. Area Frame Sampling in Agriculture. SRS-20. U.S. Dept. Agr., Stat. Rept. Serv., 1975.

(6) Huddleston, H.F., P.L. Claypool, and R.R. Hocking. Optimal Allocation to Strata Using Convex Programming. The Journal of the Royal Statistical Society, Series C, Vol. 19, No. 3, 1970, pp. 273-278.

(7) Pratt, William L. The Use of Interpenetrating Sampling in Area Frames. U.S. Dept. Agr., Stat. Rept. Serv., May 1974.

(8) U.S. Department of Agriculture. Scope and Methods of the Statistical Reporting Service, Stat. Rept. Serv., Publ. No. 1308, July 1975, p.8.

(9) U.S. Geological Survey. LANDSAT Data Users Handbook, 1979.

A P P E N D I C E S

A R E A   F R A M E   A N A L Y S I S   P A C K A G E
FOR CALIFORNIA

13:20 THURSDAY, OCTOBER 18, 1979

ERROR FILE

✓ PLANIMETERED TO EXPECTED SIZE OUTSIDE LIMITS *****
STR=3100 SEG=685 _TYPE_=1 _FREQ_=13 IDSEG=9685 PS=5 REP=1 ROTA=9 CRD=40 CO=1 EXP=369 PLAN=204 TACR=201 TTA=13 NCA=1 PST=11 NAG=188
NUMRFO=1 TRFO=13 CULT=1 V200=0 V250=0 V511=0 V512=0 V513=0 V514=0 V515=0 V516=0 V517=0 V524=0 V535=0 V540=0 V549=0 V552=0 V553=0
V570=0 V574=0 V605=0 V691=0 ESS=160 CK=0 EXSIZE=160 CK1=0 _ERROR_=1 _N_=1   ODD SHAPE - LOOKS OKAY   EXP C.U. SIZE = 190
ROUNDING OF C.U. → COUNT UNIT SIZE

** PLANIMETERED TO EXPECTED SIZE OUTSIDE LIMITS *****
STR=3100 SEG=693 _TYPE_=1 _FREQ_=3 IDSEG=9693 PS=5 REP=2 ROTA=9 CRD=40 CO=1 EXP=369 PLAN=207 TACR=211 TTA=30 NCA=0 PST=30 NAG=181
NUMRFO=1 TRFO=31 CULT=0 V200=0 V250=20 V511=0 V512=0 V513=0 V514=0 V515=0 V516=0 V517=0 V524=0 V535=0 V540=0 V549=0 V552=0 V553=0
V570=0 V574=0 V605=0 V691=0 ESS=160 CK=0 EXSIZE=160 CK1=0 _ERROR_=1 _N_=2   PI = 130   FRAME = 180   ENUMERATION OR PLANIMETERING
CHECK SEG FOR OVERENUMERATION

PLANIMETERED TO EXPECTED SIZE OUTSIDE LIMITS *****
STR=1700 SEG=245 _TYPE_=1 _FREQ_=2 IDSEG=9245 PS=5 REP=1 ROTA=9 CRD=50 CO=7 EXP=43 PLAN=232 TACR=225 TTA=165 NCA=0 PST=165 NAG=60
NUMRFO=1 TRFO=19000 CULT=0 V200=0 V250=0 V511=0 V512=0 V513=0 V514=0 V515=0 V516=0 V517=0 V524=0 V535=0 V540=0 V549=0 V552=0 V553=0
V570=0 V574=0 V605=0 V691=0 ESS=320 CK=0 EXSIZE=320 CK1=0 _ERROR_=1 _N_=3   PI = 224   COULD HAVE BEEN STRATIFIED AS 41 OR 20
STRATIFICATION

** OBSERVED TO PLANIMETERED SIZE OUTSIDE LIMITS *****
STR=4100 SEG=764 _TYPE_=1 _FREQ_=16 IDSEG=9764 PS=4 REP=4 ROTA=9 CRD=60 CO=9 EXP=104 PLAN=2450 TACR=1585 TTA=877 NCA=0 PST=872
NAG=708 NUMRFO=3 TRFO=7032 CULT=5 V200=0 V250=115 V511=0 V512=0 V513=0 V514=0 V515=0 V516=0 V517=0 V524=0 V535=0 V540=0 V549=0
V552=0 V553=0 V570=0 V574=0 V605=0 V691=0 ESS=2560 CK=0 EXSIZE=2560 CK1=0 _ERROR_=1 _N_=4   NO PI   BOUNDARIES DIFFICULT
QUAD ≐ 2200   FRAME = 2300   ENUMERATION

PUNCH OBSERVED TO PLANIMETERED SIZE OUTSIDE LIMITS *****   SEG = 824   TACR = 2811   TTA = 2669   TRFO = 2811
STR=4100 SEG=824 _TYPE_=1 _FREQ_=9 IDSEG=9824 PS=4 REP=10 ROTA=9 CRD=60 CO=9 EXP=104 PLAN=2671 TACR=342 TTA=200 NCA=0 PST=195
NAG=142 NUMRFO=4 TRFO=362 CULT=5 V200=0 V250=7 V511=0 V512=0 V513=0 V514=0 V515=0 V516=0 V517=0 V524=0 V535=0 V540=0 V549=0 V552=0
V553=0 V570=0 V574=0 V605=0 V691=0 ESS=2560 CK=0 EXSIZE=2560 CK1=5 _ERROR_=1 _N_=5 ONE LARGE PART C QUESTIONNAIRE NOT PUNCHED
keypunch

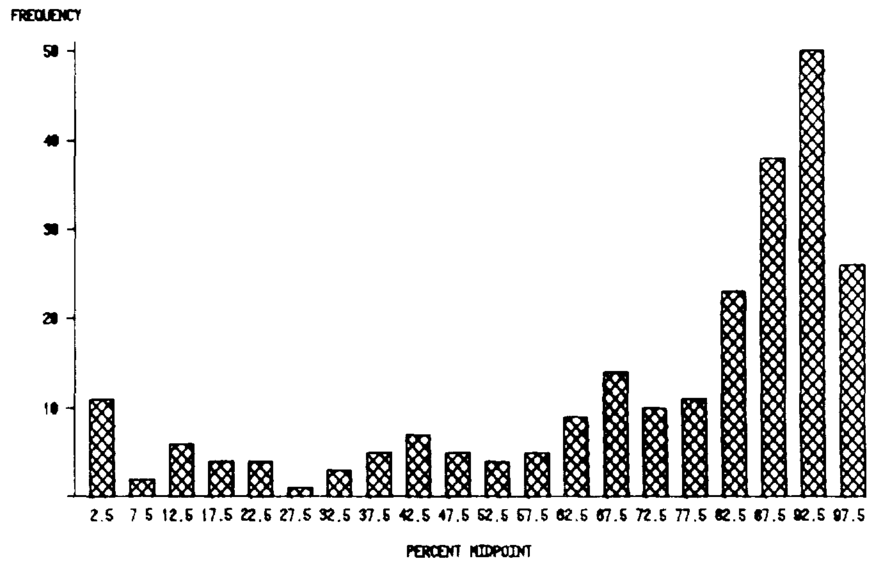
* PLANIMETERED TO EXPECTED SIZE OUTSIDE LIMITS *****
STR=1300 SEG=75 _TYPE_=1 _FREQ_=25 IDSEG=9075 PS=3 REP=4 ROTA=9 CRD=40 CO=13 EXP=29 PLAN=470 TACR=472 TTA=119 NCA=9 PST=103 NAG=353
NUMRFO=8 TRFO=122 CULT=7 V200=4 V250=36 V511=0 V512=0 V513=0 V514=0 V515=0 V516=0 V517=0 V524=0 V535=0 V540=0 V549=0 V552=0 V553=0
V570=0 V574=0 V605=0 V691=0 ESS=640 CK=0 EXSIZE=640 CK1=0 _ERROR_=1 _N_=6   PI = 480   SEGMENT NEXT TO WATER   STRATIFICATION
COULD & SHOULD BE STRATA 31

✓ OBSERVED TO PLANIMETERED SIZE OUTSIDE LIMITS *****
STR=4100 SEG=732 _TYPE_=1 _FREQ_=1 IDSEG=9732 PS=2 REP=1 ROTA=9 CRD=10 CO=15 EXP=104 PLAN=2560 TACR=2000 TTA=0 NCA=0 PST=0 NAG=2000
NUMRFO= TRFO=0 CULT=0 V200=0 V250=0 V511=0 V512=0 V513=0 V514=0 V515=0 V516=0 V517=0 V524=0 V535=0 V540=0 V549=0 V552=0 V553=0
V570=0 V574=0 V605=0 V691=0 ESS=2560 CK=0 EXSIZE=2560 CK1=5 _ERROR_=1 _N_=7   FRAME = 2000   N.P.

✓ OBSERVED TO PLANIMETERED SIZE OUTSIDE LIMITS *****
STR=4100 SEG=782 _TYPE_=1 _FREQ_=7 IDSEG=9782 PS=2 REP=6 ROTA=9 CRD=10 CO=15 EXP=104 PLAN=2560 TACR=2240 TTA=0 NCA=0 PST=0 NAG=2240
NUMRFO= TRFO=0 CULT=0 V200=0 V250=0 V511=0 V512=0 V513=0 V514=0 V515=0 V516=0 V517=0 V524=0 V535=0 V540=0 V549=0 V552=0 V553=0
V570=0 V574=0 V605=0 V691=0 ESS=2560 CK=0 EXSIZE=2560 CK1=5 _ERROR_=1 _N_=8   FRAME = 2380   N.P.

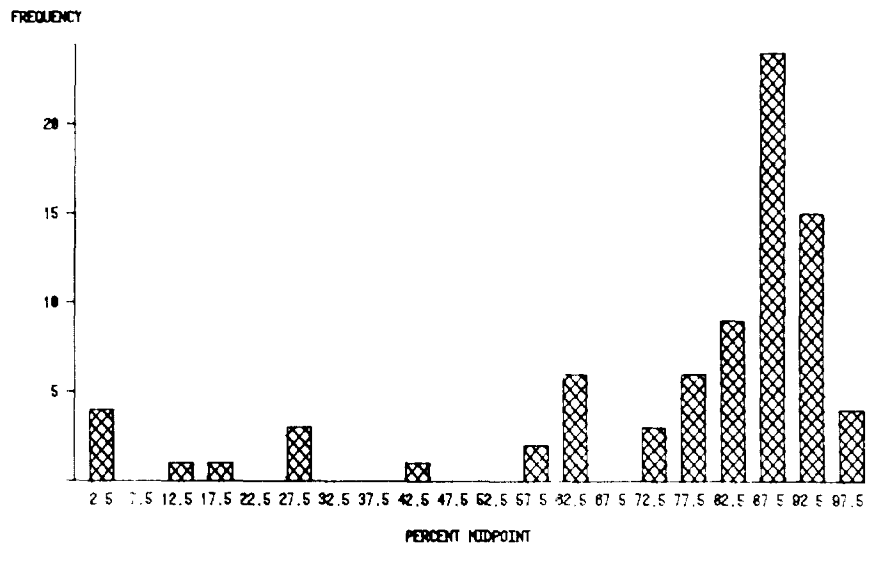
L.F. ✓ OBSERVED TO PLANIMETERED SIZE OUTSIDE LIMITS *****
STR=1300 SEG=39 _TYPE_=1 _FREQ_=2 IDSEG=9039 PS=15 REP=2 ROTA=9 CRD=51 CO=19 EXP=29 PLAN=640 TACR=161 TTA=160 NCA=5 PST=0 NAG=1
NUMRFO=1 TRFO=2600 CULT=155 V200=0 V250=0 V511=0 V512=0 V513=0 V514=0 V515=0 V516=0 V517=0 V524=155 V535=0 V540=0 V549=0 V552=0
V553=0 V570=0 V574=0 V605=0 V691=0 ESS=640 CK=0 EXSIZE=640 CK1=5 _ERROR_=1 _N_=9   PI = 640   N.P.

APPENDIX FIGURE 1--NUMBER OF SEGMENTS BY PERCENT CULTIVATED LAND
FOR STRATUM 17



NOTE: EXCLUDES SEGMENTS CONTAINING LARGE FARM TRACTS

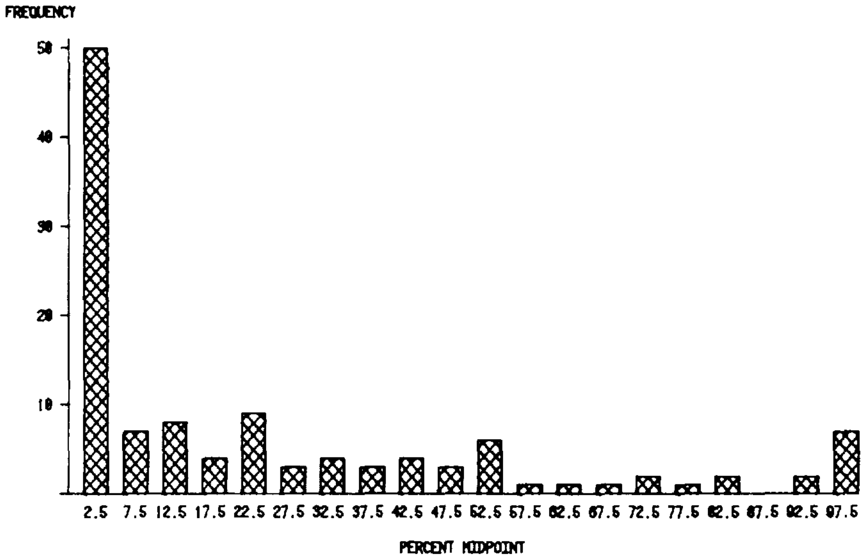
APPENDIX FIGURE 2--NUMBER OF SEGMENTS BY PERCENT CULTIVATED LAND
FOR STRATUM 19



NOTE: EXCLUDES SEGMENTS CONTAINING LARGE FARM TRACTS

APPENDIX FIGURE 3--NUMBER OF SEGMENTS BY PERCENT CULTIVATED LAND

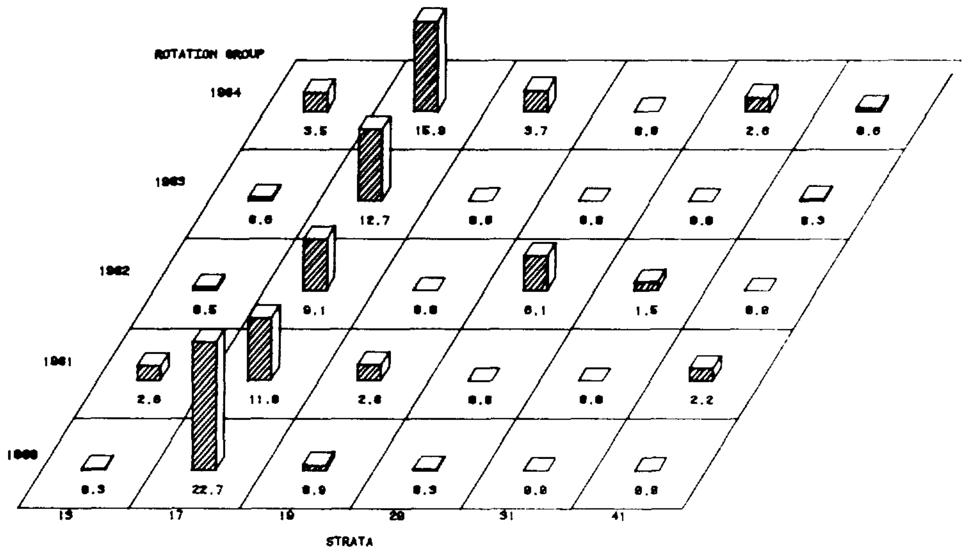
FOR STRATUM 20



NOTE: EXCLUDES SEGMENTS CONTAINING LARGE FARM TRACTS

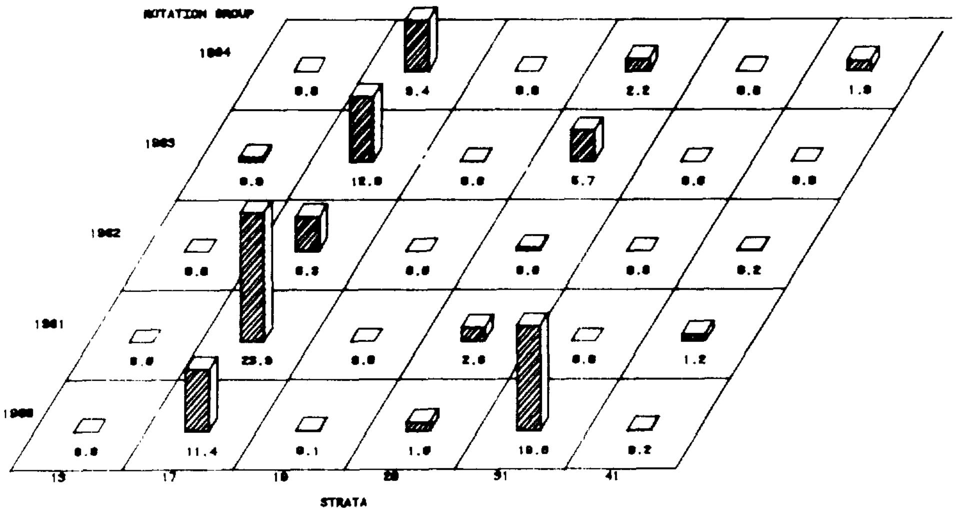
APPENDIX FIGURE 4--PERCENTAGE OF ESTIMATED GRAPES (ALL TYPES) BY ROTATION GROUP FOR SELECTED STRATA, 1979

BLOCK CHART OF PCTGRPS



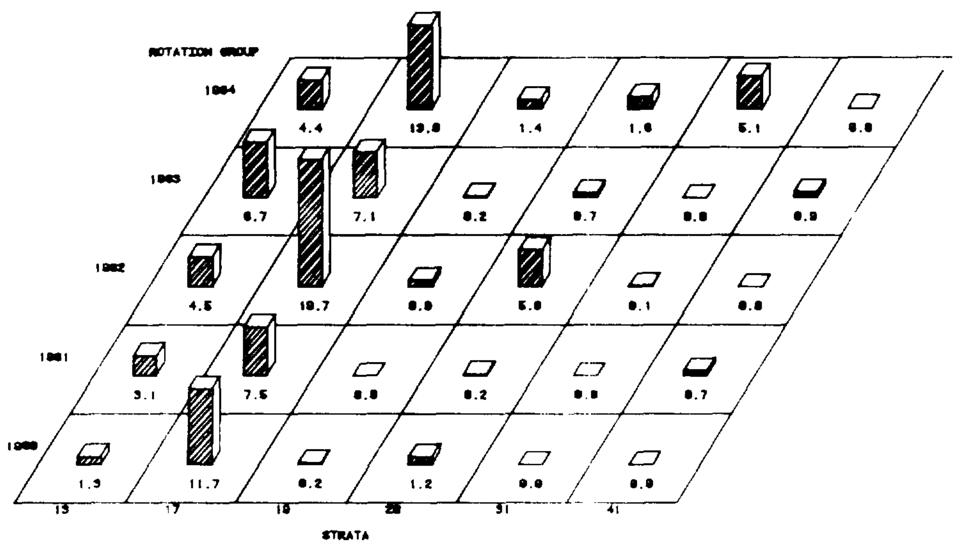
APPENDIX FIGURE 5--PERCENTAGE OF ESTIMATED ORANGES (ALL) BY ROTATION GROUP FOR SELECTED STRATA, 1979

BLOCK CHART OF PCTORNG

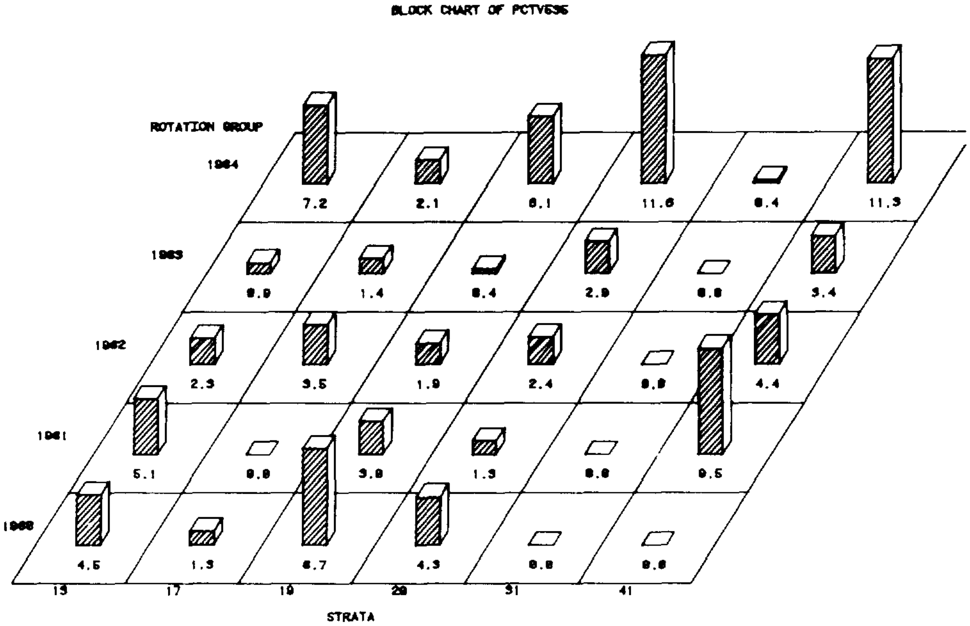


APPENDIX FIGURE 6--PERCENTAGE OF ESTIMATED TREE NUTS BY ROTATION GROUP FOR SELECTED STRATA, 1979

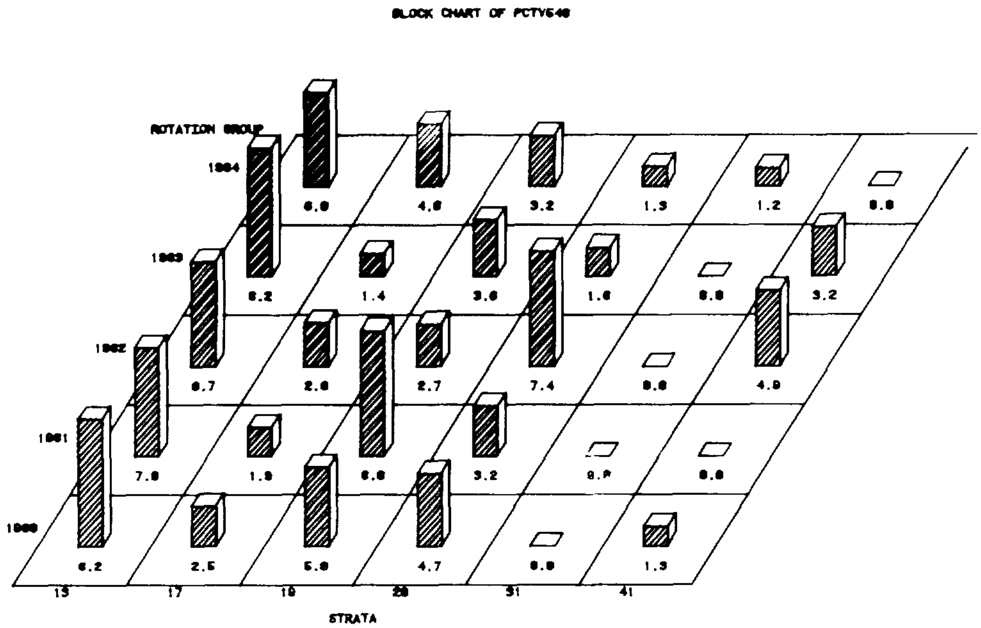
BLOCK CHART OF PCTNUTS



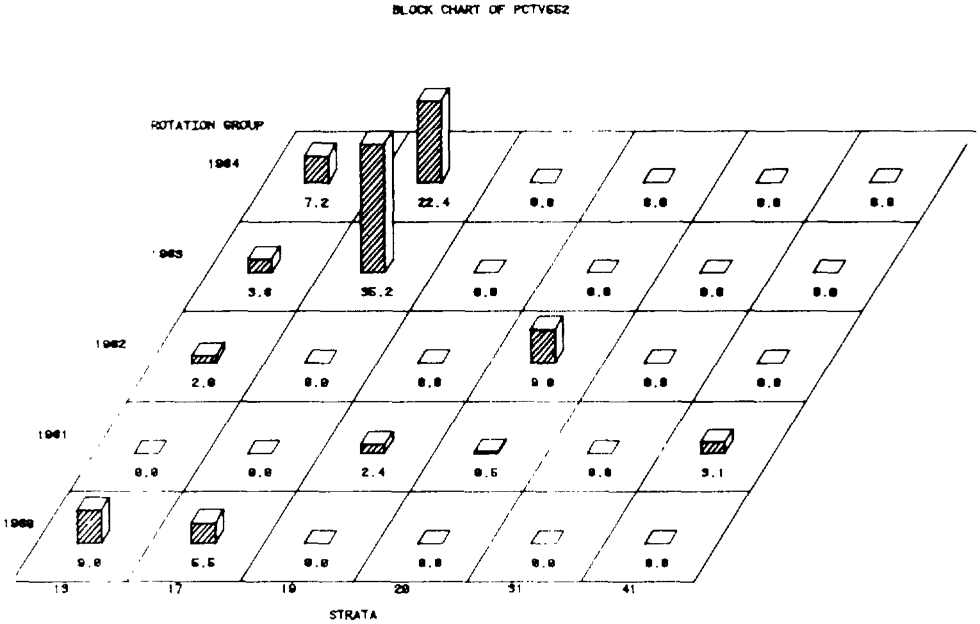
APPENDIX FIGURE 7--PERCENTAGE OF ESTIMATED BARLEY BY ROTATION GROUP FOR SELECTED STRATA, 1979



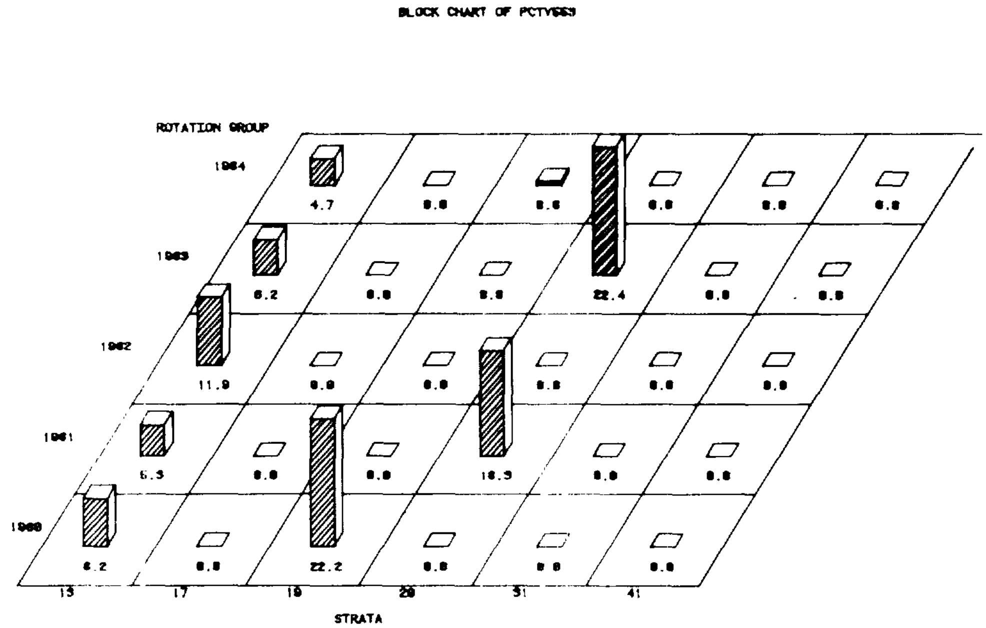
APPENDIX FIGURE 8--PERCENTAGE OF ESTIMATED WINTER WHEAT BY ROTATION GROUP FOR SELECTED STRATA, 1979



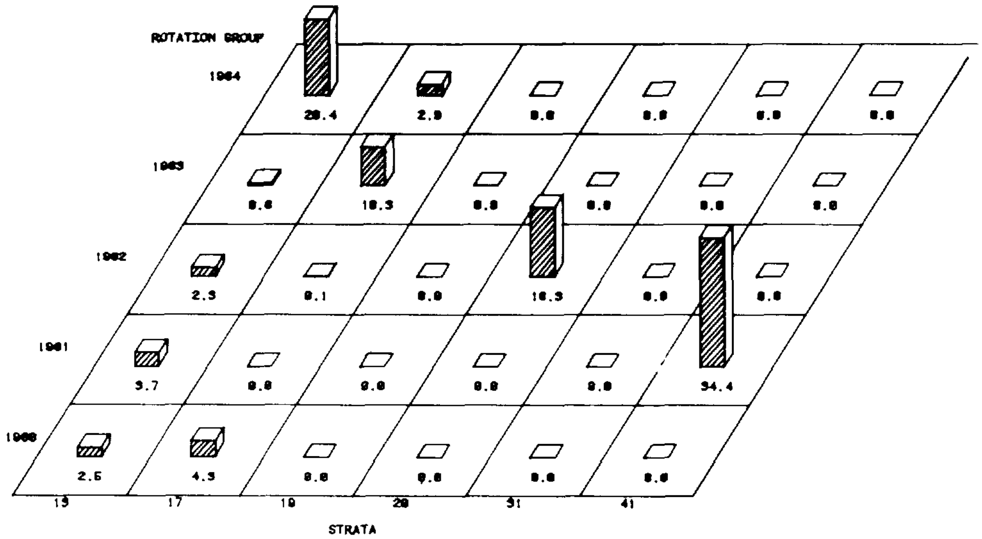
APPENDIX FIGURE 9--PERCENTAGE OF ESTIMATED POTATOES BY ROTATION GROUP FOR SELECTED STRATA, 1979



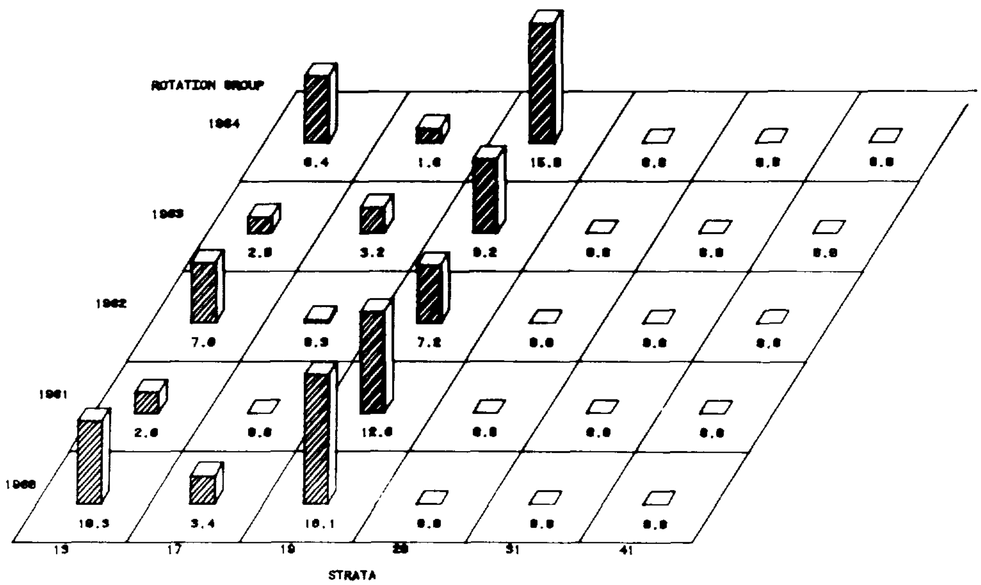
APPENDIX FIGURE 10--PERCENTAGE OF ESTIMATED DURUM WHEAT BY ROTATION GROUP FOR SELECTED STRATA, 1979



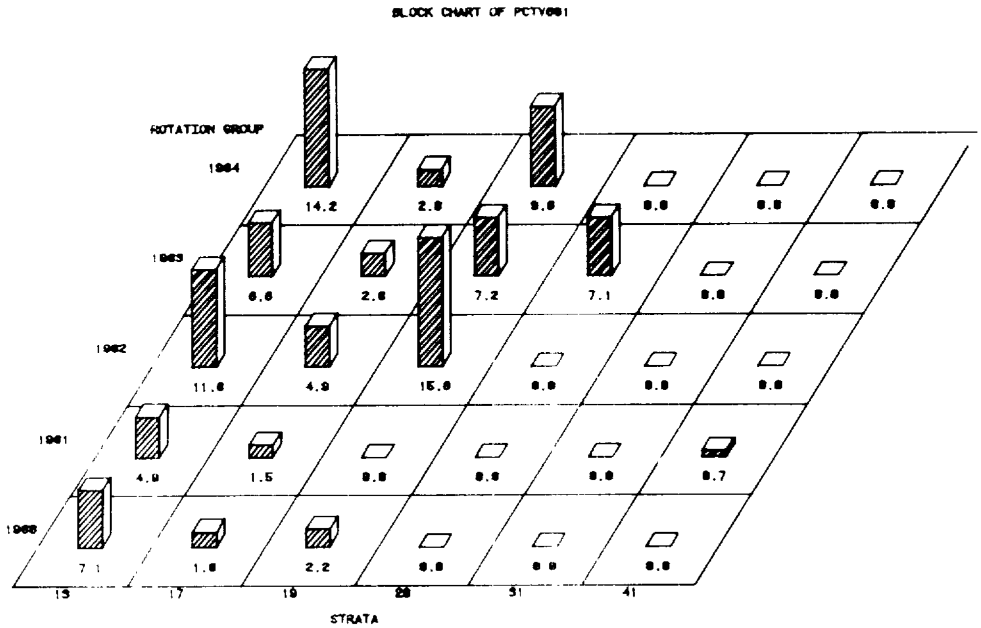
APPENDIX FIGURE 11--PERCENTAGE OF ESTIMATED SORGHUM BY ROTATION GROUP FOR SELECTED STRATA, 1979



APPENDIX FIGURE 12--PERCENTAGE OF ESTIMATED TOMATOES BY ROTATION GROUP FOR SELECTED STRATA, 1979



APPENDIX FIGURE 13--PERCENTAGE OF ESTIMATED SUGAR BEETS BY ROTATION GROUP FOR SELECTED STRATA, 1979



A R E A   F R A M E   A N A L Y S I S   P A C K A G E
FOR CALIFORNIA
MODULE NAME: TWOVAR
PAPER STRATA ESTIMATES FOR COTTON
V524PST=ESTIMATED TOTALS
STRATA=STRATA   PS=PAPER STRATA
PSSE=STANDARD ERROR   PSCV=COEFFICIENT OF VARIATION
MV524=MAXIMUM SEGMENT EXPANSION
AVGEXP=AVERAGE EXPANSION FACTOR

OBS	STRATA	PS	_FREQ_	V524PST	PSSE	PSCV	MV524	V524RNGE	AVGEXP
1	1300	1	10	0	0.0	0.0000	0	0	29
2	1300	2	10	0	0.0	0.0000	0	0	29
3	1300	3	10	0	0.0	0.0000	0	0	29
4	1300	4	10	0	0.0	0.0000	0	0	29
5	1300	5	10	0	0.0	0.0000	0	0	29
6	1300	6	10	0	0.0	0.0000	0	0	29
7	1300	7	10	0	0.0	0.0000	0	0	29
8	1300	8	10	0	0.0	0.0000	0	0	29
9	1300	9	10	0	0.0	0.0000	0	0	29
10	1300	10	10	0	0.0	0.0000	0	0	29
11	1300	11	10	0	0.0	0.0000	0	0	29
12	1300	12	10	0	0.0	0.0000	0	0	29
13	1300	13	10	0	0.0	0.0000	0	0	29
14	1300	14	10	65047	13810.6	21.2317	12325	12325	29
15	1300	15	10	60668	18624.3	31.0283	16733	16733	29
16	1300	16	10	61915	10542.3	17.0270	11716	11716	29
17	1300	17	10	73370	9536.5	12.9977	11803	8729	29
18	1300	18	10	48662	20067.0	41.2375	17545	17545	29
19	1300	19	10	50779	19047.9	37.5114	17980	17980	29
20	1300	20	10	83085	14283.5	17.1914	16240	12354	29
21	1300	21	10	70731	18590.2	26.2829	18879	18879	29
22	1300	22	10	62756	18181.3	28.9713	19314	19314	29
23	1300	23	10	69624	22638.2	32.5097	22069	22069	29
24	1300	24	10	17920	9955.5	55.5548	9940	9940	28
25	1700	1	10	0	0.0	0.0000	0	0	43
26	1700	2	10	0	0.0	0.0000	0	0	43
27	1700	3	10	0	0.0	0.0000	0	0	43
28	1700	4	10	0	0.0	0.0000	0	0	43
29	1700	5	10	0	0.0	0.0000	0	0	43
30	1700	6	10	0	0.0	0.0000	0	0	43
31	1700	7	10	0	0.0	0.0000	0	0	43
32	1700	8	10	0	0.0	0.0000	0	0	43
33	1700	9	10	0	0.0	0.0000	0	0	43
34	1700	10	10	0	0.0	0.0000	0	0	43
35	1700	11	10	0	0.0	0.0000	0	0	43
36	1700	12	10	13670	6688.2	48.9113	5848	5848	43
37	1700	13	10	34615	10125.7	29.2522	9417	9417	43
38	1700	14	10	2830	1321.6	46.5660	1204	1204	43
39	1700	15	10	21202	11710.2	55.1272	9288	9288	43
40	1700	16	10	18708	9482.2	50.5767	6880	6880	43
41	1700	17	10	29552	10037.9	33.1565	8170	8170	43
42	1700	18	10	14534	8297.2	57.0880	7912	7912	43
43	1700	19	10	20597	10426.3	50.6201	9890	9890	43
44	1700	20	10	34572	16009.4	46.3047	13244	13244	43
45	1700	21	10	38141	12074.1	31.6565	10320	10320	43
46	1700	22	10	55126	17069.4	30.9642	12400	12400	43
47	1700	23	10	21554	8445.0	39.1407	6880	6880	43
48	1700	24	10	7098	5114.0	72.0477	4956	4956	42
49	1900	1	8	0	0.0	0.0000	0	0	45
50	1900	2	8	0	0.0	0.0000	0	0	45
51	1900	3	8	0	0.0	0.0000	0	0	45
52	1900	4	8	0	0.0	0.0000	0	0	45
53	1900	5	8	0	0.0	0.0000	0	0	45
54	1900	6	8	51400	23572.5	46.7704	23670	23670	45
55	1900	7	8	104445	20942.6	20.0322	21375	21375	45
56	1900	8	8	97110	23943.1	24.6556	27765	27765	45
57	1900	9	8	25425	7626.0	29.9938	7065	7065	45
58	1900	10	8	9900	6718.9	67.8673	6300	6300	45
59	2000	1	10	0	0.0	0.0000	0	0	64
60	2000	2	10	0	0.0	0.0000	0	0	64
61	2000	3	10	0	0.0	0.0000	0	0	64
62	2000	4	10	0	0.0	0.0000	0	0	64
63	2000	5	10	0	0.0	0.0000	0	0	64
64	2000	6	10	0	0.0	0.0000	0	0	64
65	2000	7	10	0	0.0	0.0000	0	0	64
66	2000	8	10	0	0.0	0.0000	0	0	64
67	2000	9	10	24576	24383.2	99.2153	24576	24576	64
68	2000	10	10	42432	38182.9	89.9859	38720	38720	64
69	2000	11	10	47552	22300.2	46.8963	19840	19840	64
70	2000	12	10	0	0.0	0.0000	0	0	65
71	3100	1	5	0	0.0	0.0000	0	0	369
72	3100	2	5	0	0.0	0.0000	0	0	369
73	3100	3	5	0	0.0	0.0000	0	0	369
74	3100	4	5	0	0.0	0.0000	0	0	369
75	3100	5	5	0	0.0	0.0000	0	0	369
76	3100	6	5	0	0.0	0.0000	0	0	369
77	3100	7	5	16605	16582.5	99.8638	16605	16605	369
78	3100	8	5	0	0.0	0.0000	0	0	370
79	3200	1	5	0	0.0	0.0000	0	0	2315
80	3200	2	5	0	0.0	0.0000	0	0	2315
81	4100	1	10	0	0.0	0.0000	0	0	104
82	4100	2	10	0	0.0	0.0000	0	0	104
83	4100	3	10	0	0.0	0.0000	0	0	104
84	4100	4	10	0	0.0	0.0000	0	0	104
85	4100	5	10	0	0.0	0.0000	0	0	104
86	4100	6	10	0	0.0	0.0000	0	0	104
87	4100	7	10	0	0.0	0.0000	0	0	104
88	4100	8	10	0	0.0	0.0000	0	0	104
89	4100	9	10	0	0.0	0.0000	0	0	104
90	4100	10	10	96928	80793.0	83.3535	81432	81432	104
91	4300	1	4	0	0.0	0.0000	0	0	199
92	4300	2	4	0	0.0	0.0000	0	0	199
93	4300	3	4	0	0	0	0	0	199.0
94	4300	4	4	0	0	0	0	0	199.0
95	4300	5	4	0	0	0	0	0	199.0
96	4400	1	5	0	0	0	0	0	112.8
97	4400	2	5	0	0	0	0	0	73.8
98	4400	3	5	0	0	0	0	0	39.8
99	4400	4	5	0	0	0	0	0	1[illegible].0
100	.	.		0	.	.	0	0	57.[illegible]
101	4500	1	4	0	0	0	0	0	501.0
102	4500	2	4	0	0	0	0	0	501.0
103	5000	1	4	0	0	0	0	0	886.0
104	5000	2	4	0	0	0	0	0	886.0

AREA FRAME ANALYSIS PACKAGE
FOR CALIFORNIA
MODULE NAME: TWOVAR
STATE AND STRATA ESTIMATES FOR COTTON
STRATA=STRATA _FREQ_=NUMBER OF PAPER STRATA
V524STRT=ESTIMATED TOTAL
STRSE=STANDARD ERROR STRCV=COEFFICIENT OF VARIATION
PCTEST=STRATUM PERCENTAGE OF TOTAL ESTIMATE
PCTVAR=STRATUM PERCENTAGE OF TOTAL VARIANCE

OBS	LEVEL	STRATA	_FREQ_	V524STRT	STRSE	STRCV	PCTEST	PCTVAR
1	STATE	.	104	1490226	124360	8.3450	100.000	100.000
2		1300	24	664562	54812	8.2478	44.595	19.426
3		1700	24	310291	37930	12.2241	20.822	9.303
4		1900	10	287280	40865	14.2249	19.278	10.798
5		2000	12	114560	50495	44.0776	7.687	16.487
6		3100	[illegible]	16605	16582	99.8658	1.114	1.778
7		3200	[illegible]	0	0	0.0000	0.000	0.000
8		4100	10	95928	80793	83.5535	6.504	42.207
9		4300	5	0	0	0.0000	0.000	0.000
10		4400	5	0	0	0.0000	0.000	0.000
11		4500	2	0	0	0.0000	0.000	0.000
12		5000	2	0	0	0.0000	0.000	0.000

AREA FRAME ANALYSIS PACKAGE
FOR CALIFORNIA
MODULE NAME: TWOVAR
REPLICATE ESTIMATE
STRATA=STRATA REP=REPLICATION
REPN=NUMBER OF REPLICATES IN THE STRATA
PSN=PAPER STRATA SIZE
V524REPT=ESTIMATED TOTAL
MV524=MAXIMUM SEGMENT EXPANSION

OBS	STRATA	REP	REPN	PSN	V524REPT	MV524
1	1300	1	10	24	799530	22069
2	1300	2	10	24	672860	19514
3	1300	3	10	24	913500	18879
4	1300	4	10	24	503730	16733
5	1300	5	10	24	659460	11107
6	1300	6	10	24	720760	17545
7	1300	7	10	24	718290	12209
8	1300	8	10	24	428620	10672
9	1300	9	10	24	562020	12702
10	1300	10	10	24	602050	16240
11	1700	1	10	24	296700	13244
12	1700	2	10	24	82560	4902
13	1700	3	10	24	260260	6622
14	1700	4	10	24	365930	11481
15	1700	5	10	24	554620	10148
16	1700	6	10	24	407640	9116
17	1700	7	10	24	247680	10320
18	1700	8	10	24	376680	12513
19	1700	9	10	24	286380	8170
20	1700	10	10	24	224460	12900
21	1900	1	8	10	262800	12555
22	1900	2	8	10	306000	15975
23	1900	3	8	10	13690	1710
24	1900	4	8	10	537840	27765
25	1900	5	8	10	146520	6430
26	1900	6	8	10	306000	20025
27	1900	7	8	10	446040	23670
28	1900	8	8	10	279360	17145
29	2000	1	10	12	0	0
30	2000	2	10	12	0	0
31	2000	3	10	12	51200	5120
32	2000	4	10	12	517120	38720
33	2000	5	10	12	0	0
34	2000	6	10	12	0	0
35	2000	7	10	12	96000	9600
36	2000	8	10	12	0	0
37	2000	9	10	12	37120	3712
38	2000	10	10	12	440160	24576
39	3100	1	5	8	0	0
40	3100	2	5	8	0	0
41	3100	3	5	8	83025	16605
42	3100	4	5	8	0	0
43	3100	5	5	8	0	0
44	3200	1	5	2	0	0
45	3200	2	5	2	0	0
46	3200	3	5	2	0	0
47	3200	4	5	2	0	0
48	3200	5	5	2	0	0
49	4100	1	10	10	0	0
50	4100	2	10	10	0	0
51	4100	3	10	10	154960	15496
52	4100	4	10	10	0	0
53	4100	5	10	10	0	0
54	4100	6	10	10	814320	81432
55	4100	7	10	10	0	0
56	4100	8	10	10	0	0
57	4100	9	10	10	0	0
58	4100	10	10	10	0	0
59	4300	1	4	5	0	0
60	4300	2	4	5	0	0
61	4300	3	4	5	0	0
62	4300	4	4	5	0	0
63	4400	1	5	5	0	0
64	4400	2	5	5	0	0
65	4400	3	5	5	0	0
66	4400	4	5	5	0	0
67	4400	5	5	5	0	0
68	4500	1	4	2	0	0
69	4500	2	4	2	0	0
70	4500	3	4	2	0	0
71	4500	4	4	2	0	0
72	5000	1	4	2	0	0
73	5000	2	4	2	0	0
74	5000	3	4	2	0	0
75	5000	4	4	2	0	0

AREA FRAME ANALYSIS PACKAGE
FOR CALIFORNIA
MODULE NAME: TWOVAR
LAND USE STRATA REPLICATE TOTAL ESTIMATE FOR COTTON
STRATA=STRATA   V524STRT=STRATA TOTAL
CV=COEFFICIENT OF VARIATION OF THE TOTAL ESTIMATOR
V524STSE=STRATA STANDARD ERROR

OBS	STRATA	_FREQ_	V524STRT	CV	V524STSE
1	1300	10	664562	6.963	46271.9
2	1700	10	310291	12.850	34872.7
3	1900	8	287280	19.954	57324.0
4	2000	10	114560	54.175	62063.3
5	3100	5	16605	99.999	16605.0
6	3200	5	0	0.000	0.0
7	4100	10	95928	83.757	81184.2
8	4300	4	0	0.000	0.0
9	4400	5	0	0.000	0.0
10	4500	4	0	0.000	0.0
11	5000	4	0	0.000	0.0

AREA FRAME ANALYSIS PACKAGE
FOR CALIFORNIA
MODULE NAME: TWOVAR
STATE REPLICATED ESTIMATE FOR COTTON
V524TL=STATE TOTAL
CVV=COEFFICIENT OF VARIATION OF THE STATE TOTAL ESTIMATOR
SERR=STANDARD ERROR OF THE ESTIMATOR

OBS	V524TL	CVV	SERR
1	1490226	8.93652	133174

*U.S. GOVERNMENT PRINTING OFFICE : [illegible] -140-012/ESS-188