EVALUATION OF PROCEDURES FOR ESTIMATING CITRUS FRUIT YIELD

FRUIT
COUNTS

GROUND
PHOTOGRAPHY

REMOTE
SENSING

STATISTICAL REPORTING SERVICE - U.S. DEPARTMENT OF AGRICULTURE

CONTENTS

Washington, D. C. 20250 February 1972

SUMMARY AND CONCLUSIONS

Color photography and limb counts can be used to forecast citrus fruit production, according to USDA-sponsored research in Texas in 1969-70. Estimates of variances and costs of these and other objective yield procedures were obtained in a random sample of citrus blocks in the Lower Rio Grande Valley as a followup of preliminary research in 1968. Findings based on fruit size and droppage through the 1969 growing season included the following:

(1) A sample of 65-85 blocks each of early oranges, Valencia oranges, and grapefruit is sufficient to estimate fruit set early in the growing season with a coefficient of variation of 10 percent. Optimum allocations for counting fruit on sample limbs are two trees per block, two primary limbs per tree, and two terminal limbs per primary.

(2) Correlation of estimated fruit per tree and photo count was high for all types of citrus fruit studied. [Relationship of estimated fruit per tree to fruit count from photos did not vary significantly among blocks.]

(3) Time required to identify terminal (count) limbs was reduced with a newly devised limb selection gauge.

(4) Seasonal fruit growth patterns differed somewhat from the 1968 growing season. Differences in fruit set between the two seasons affected average size in some blocks of fruit.

(5) Amount of fruit droppage did not differ significantly between the 1968 and 1969 growing seasons.

(6) Analyses of fruit samples for quality and expected size distribution at maturity may prove as valuable as information on expected yield.

EVALUATION OF PROCEDURES FOR ESTIMATING CITRUS FRUIT YIELD

Fruit Counts, Ground Photography, Remote Sensing

by
Richard D. Allen
Standards and Research Division
Statistical Reporting Service

BACKGROUND

In 1968, the Research and Development Branch, Statistical Reporting Service (SRS), U.S. Department of Agriculture (USDA), conducted a study of ground photography and other objective procedures to forecast yields of oranges and grapefruit in selected groves in the Lower Rio Grande Valley. Texas Citrus Mutual and the Remote Sensing Laboratory, Agricultural Research Service (ARS), USDA, Weslaco, Texas, cooperated in the study. That research indicated the feasibility of using a variety of new techniques--including photography--to arrive at objective yield estimates. 1/

During the 1969 growing season, a followup study was made in an attempt to resolve some of the problems raised in the earlier survey and to further refine fruit counting procedures. Specific objectives of the project included the following:

(1) Improvement of sampling definitions and procedures for objective yield surveys.

(2) Estimation of components of variance for the various sources of variation in estimating number of citrus fruit per tree.

(3) Collection of cost information for various objective yield procedures.

(4) Evaluation of the use of ground photography in a sampling design with estimates of total fruit rather than actual counts.

(5) Study of ways aerial photography might be used to improve estimation techniques.

1/ Estimation based on actual plant or fruit characteristics measured or counted from randomly selected plots or limbs.

The Research and Development Branch, SRS, was responsible for conducting the 1969 research study. Field work was carried out by the R&D Branch, Texas State Statistical Office of SRS, and Texas Citrus Mutual. Financing was shared by SRS, Texas Citrus Mutual, and the Texas Department of Agriculture. The ARS Remote Sensing Laboratory at Weslaco, Tex., assisted with some portions of the research.

SAMPLE SELECTION

A listing of citrus trees compiled by Texas Citrus Mutual in the Lower Rio Grande Valley was used as the sampling frame. Trees were designated by type of citrus (early oranges, Valencia oranges, or grapefruit) and age (0-3 years, 4-7 years, or 8 years and over) at the time of listing (about 1967). A separate sample was selected for each age class of the three types.

All orchards were listed by irrigation district, subdistrict, block, and lot number. Another code was used to indicate the owner of each grove of fruit.

For each type-age combination, a systematic random sample was drawn, giving each grove in that listing the same chance of selection. Thus, all sizes of groves had the same probability of selection. Desired sample sizes for the field work were three groves 0-3 years old, nine groves 4-7 years old, and nine groves 8 years old and over. Ten groves were selected from each 0-3-year-old listing and 20 groves from each of the 4-7 and 8-years-and-over listings. This larger sample was selected to provide replacements for groves that had been removed since the listing was compiled. Since the systematic samples from the sampling frame were in sequence, the groves selected were placed in random order. Otherwise, all groves included for field work would have been located in only one county.

To cut costs, the Valencia orange workload was reduced. The 8-years-and-over sample groves were not visited, but the desired numbers of groves were worked for the other two age classes.

In addition to the random sample of groves selected for the variance study, research continued in the eight blocks of fruit that had been studied in the 1968 growing season. These blocks were used for training field crews in counting procedures. Fruit on selected limbs in these blocks were tagged for comparison of growth and droppage with the previous season.

PRESURVEY PROCEDURES

A visit was made to each selected grove before the fruit counting survey to obtain permission to make counts and measurements, to draw a field sketch of the grove, and to measure cross-sectional trunk area of a sample of trees. Trunk measurements were arrayed by size, and a systematic sample of four trees was selected for limb counts and photography. Use of trunk measurements in the selection process was intended to assure some variation in tree characteristics within a grove and to allow testing of the possible use of trunk size as an auxiliary variable in a double sampling design.

An explantory letter about the research project was prepared by Texas Citrus Mutual and sent to the owner or manager of each sample grove. No problems were encountered in obtaining permission for the survey except for the difficulty of contacting some operators. Many owners were not local residents, and management of several groves had changed at least once since the listing was originally prepared.

Field sketches were made on computer printout paper plotted with dots or small circles. Each dot or circle could be interpreted as a tree and grove boundaries indicated. Symbols such as X, G, and O were used to indicate missing trees, grapefruit trees interplanted, and orange trees interplanted.

A sample of 30 to 40 trees was desired in each grove for trunk measurements. Sampling rates of rows and trees per row depended upon total number of trees and shape of the block.

Enumerators attempted to locate all trees of designated type and age within the selected groves. Consequently, only a portion of the total grove was sketched if the different types of citrus were planted in individual rows. In some small blocks interplanted with several types of citrus, all trees had to be identified.

FIELD PROCEDURES

In the 1969 research, a two-stage limb mapping and selection process was used. Each tree was divided into first-stage units for the selection of primary limbs. Selected primaries were then divided into terminal or count limbs (see fig. 1). A random selection was made of two or more terminal limbs on which to count fruit.

To provide some control of variation in primary limb size and to ensure that each tree had a number of primary sampling units, two rules for maximum limb size were used. Cross-section area (CSA) of the trunk was measured and recorded. No primary limb could exceed one-fifth of the trunk CSA. Second, the CSA measurements of all primaries under the first rule were added. If the total primary CSA was less than trunk CSA, one-fifth of the total primary CSA was calculated; no primary was allowed to exceed this size. Most primary limbs were branches at the first scaffolding of the tree.

The one-fifth rule for determining maximum primary size did not apply to small trees (trunk CSA of 20 square inches or less). For small trees, the minimum-size requirement of at least two terminal limbs was more important, since some trees had fewer than five primary limbs. Each primary limb was given an identifying letter. During most of the 1969 work, each primary was tagged with red flagging tape. Limb letter was written on the flagging tape for later identification.

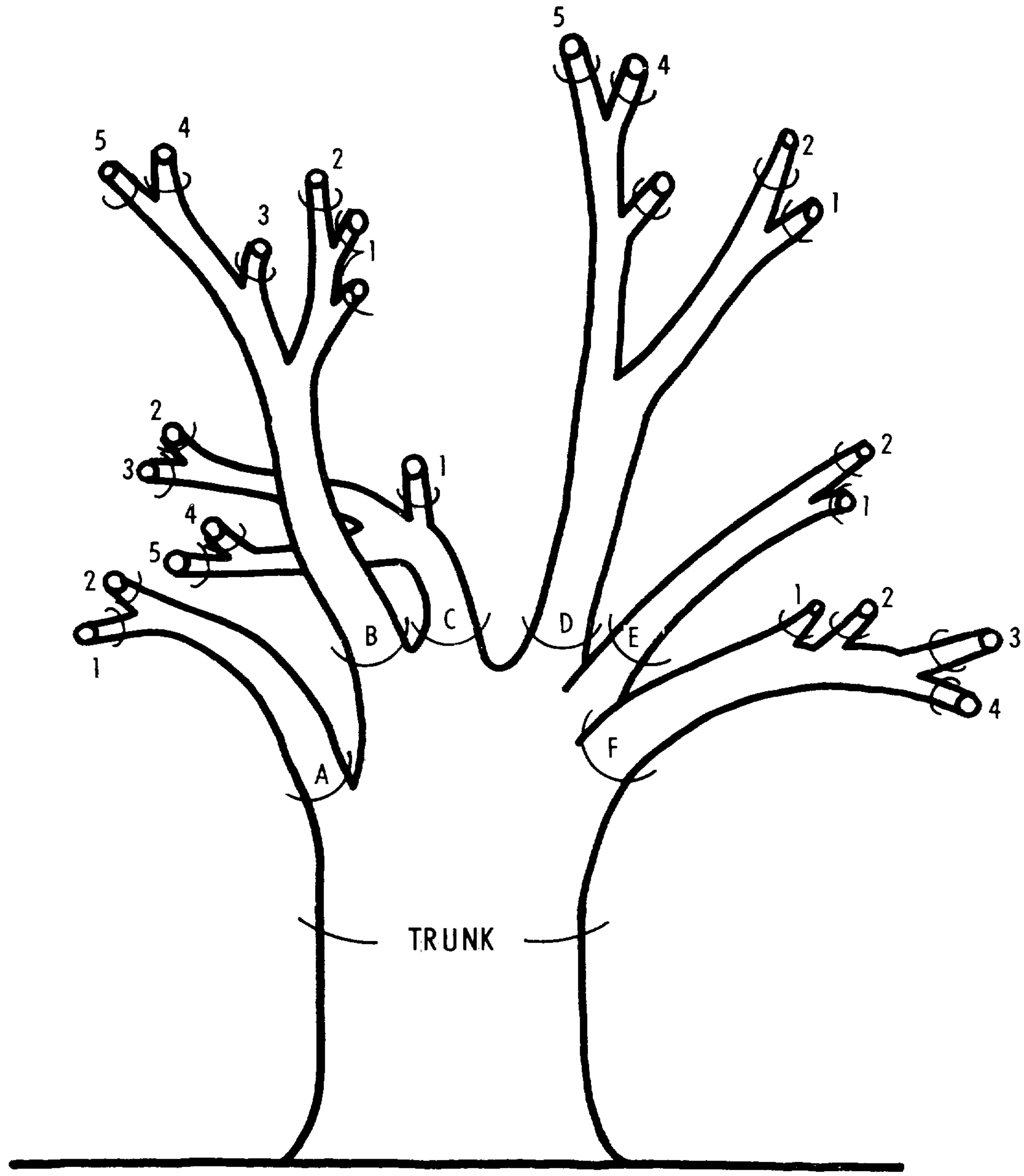

Figure 1

First- and second-stage limbs were chosen on an equal probability basis. A minimum of two primaries was selected. For early oranges and grapefruit trees, a third primary was chosen if total primary CSA was between 50 and 70 square inches. A fourth primary was selected for trees with CSA exceeding 70 square inches. Two primaries were selected on most Valencia trees, with a third primary designated on some of the larger ones. The first primary was selected by drawing a random number between one and the total number of primary limbs. Then the process was repeated. If the same number was drawn again, it was excluded and additional selections made until the desired number of primaries was selected. Thus, sampling was without replacement and number of units selected was roughly proportional to tree size.

A terminal limb was defined as a limb with a CSA larger than or equal to 0.6 square inch and smaller than or equal to 1.2 square inches. Smaller limbs (0.3, 0.4, or 0.5 square inch) could be combined with one other limb of similar or larger size to form a terminal limb if the combined total CSA did not exceed 1.2 square inches.

To facilitate mapping, a limb selection form was devised to classify the natural branching of the tree into stages. If the primary limb split into branches larger than terminal size, the CSA of each was recorded in the second-stage column, and each branch further subdivided (into third stage, fourth stage, etc.) until terminal limb size was reached. After a primary had been divided into terminal limbs, two or more terminal limbs were selected for counting. The random selection was again made on an equal probability basis.

Initially, an effort was made to determine "zero" terminals (limbs without fruit) before selection of count limbs. But this procedure was difficult and time consuming and introduced a possible source of enumerator error. Consequently, only dead limbs of terminal size were excluded before selecting terminals.

Fruit were selected for size measurements on two of the four sample trees in each grove. The sizing limbs were selected randomly from the count limbs.

All fruit on the selected limbs were tagged and numbered. Size measurements were made with circumferential calipers around the middle of the fruit at right angles to the stem.

Some trees in the variance study were revisited for additional size measurements in later survey periods. Monthly sizing visits were made to the research groves first studied in the 1968 growing season. All fruit were harvested from the sample trees in these blocks. An effort was made to time the harvest visit as close as possible to actual harvest of the grove. No harvesting was done in the random sample of blocks added in the 1969 growing season.

Nearly every tree having a limb count was photographed. Some trees may not have been photographed if they had little or no fruit. Photographs were taken from only one side of the tree. The south side was photographed if rows were east and west; in other blocks, either the west or east side was photographed, depending on the time of day.

FRUIT COUNTS

Two members of a field crew made the fruit counts generally by counting together on the same limb. Fruit were counted from the base of a limb, with each fruit being touched by one member of the team as its number was called out. One member kept track of counting progress by sliding his hand outward on the limb as it was counted.

Although this procedure appears satisfactory, it does not give an independent count. Many citrus branches are intertwined and it is often necessary for one person to hold a branch or branches away from the terminal being counted. Some terminals are quite long and it is difficult for one person to move about in the dense foliage and count all fruit on the limb. For these reasons, two people may do a better job counting together than if counts are made separately and compared. This procedure requires each member to be certain that the count is correct and that he has seen all the fruit.

Dividing the number of fruit tagged for measuring by the number originally counted gave an indication of counting accuracy. Tagging of fruit was done shortly after counting so there was little chance for fruit droppage.

Table 1 lists the quality control indications of counting accuracy by type of citrus and age of tree.

Data are based on individual limb percentages. On a total fruit basis, percentage indications were 100.00, 100.00, and 102.30 for the respective early orange classes; 101.69 and 98.42 for Valencia oranges; and 100.00, 104.81, and 106.07 for grapefruit.

Counting accuracy was within reasonable tolerances for all but the 8-and-over age categories. Previous research on citrus and other fruit indicates that counts within 2 percent are excellent. Small differences of one or two fruit on several limbs accounted for a lower percentage of limbs counted correctly (67-76 percent for four classes) than expected.

Table 1.--Quality control indications of counting accuracy, selected citrus blocks, Texas, 1969 growing season

Age class	Early oranges: Number tagged ÷ original fruit count	Early oranges: Limbs counted correctly	Valencia oranges: Number tagged ÷ original fruit count	Valencia oranges: Limbs counted correctly	Grapefruit: Number tagged ÷ original fruit count	Grapefruit: Limbs counted correctly
	Percent	Percent	Percent	Percent	Percent	Percent
0-3 years	100.00	100	100.67	92	100.00	100
4-7 years	101.46	76	99.32	85	100.75	67
8 years and over	103.73	73	1/	1/	109.20	75

1/ Valencia orange samples 8 years and over were not worked in 1969.

Expansion of Terminal Counts

The unbiased estimator of total fruit per tree from the 1969 survey is an expansion of fruit per limb by the number of limbs. Expansion factor for each terminal is the number of terminals on its primary multiplied by the total number of primary limbs. In equation form, the estimate of total fruit set from the fruit count of the ith terminal of the jth primary within a tree is expressed:

$$\hat{Y}_{ij} = m\, n_j\, x_{ij}$$

where

$\hat{Y}_{ij}$ is estimated fruit per tree,

x_{ij} is number of fruit counted,

m is the number of primary limbs on the tree,

n_j is the number of terminals on the jth primary.

The above estimate does not include path fruit. The equation with the path fruit included is

$$(\hat{Y}_{ij})_1 = m\, n_j\, x_{ij} + m\, x_{.j} + x_{..}$$

where

$x_{.j}$ is the number of path fruit on the jth primary, and

$x_{..}$ is number of path fruit associated with the trunk.

This type of estimate can be referred to as an equal probability estimate. That is, all limbs at the same stage are expanded by the same factor (the number of limbs at this stage). Relative size of the primary limb or the terminal limb of both are other possibilities for expansion. These types of expansion were also made and will be discussed later in this report.

Analysis of Expanded Data

All expanded data were identified by block number, tree number, primary letter, and terminal number within primary. Expanded data were analyzed assuming a hierarchical classification.

Fruit count estimate from the ith terminal of the jth primary of the kth tree within the lth block of trees is defined (in terms of components of variation) as follows:

$$y_{ijkl} = u + b_l + t_{kl} + p_{jkl} + e_{ijkl}.$$

Analysis of this model tests hypotheses of no differences among (1) primaries within a tree, (2) trees within a block, and (3) blocks within an age class. No attempt was made to analyze the effect of age classes at this time.

The program used for the analyses of variance also calculated the respective variance components for each level of the model. These sample estimates of variance components and estimated costs are needed in optimizing a sampling procedure.

The formula for the variance of the estimated average fruit per tree in terms of variance components is

$$\sigma^2_{\bar{y}} = \frac{\sigma_b^2}{n} + \frac{\sigma_t^2}{na} + \frac{\sigma_p^2}{nab} + \frac{\sigma_w^2}{nabc}$$

where

σ_b^2 is the block variance component,

σ_t^2 is the tree variance component,

σ_p^2 is the primary limb variance component, and

σ_w^2 is the within-primary (terminal limb) variance component.

Divisors come from the n blocks, a trees per block, b primaries per tree, and c terminals per primary that are included in the sample. Once the variance components of each level of sampling are known, the overall variance can be calculated for alternative sampling assignments.

In this sample, as in many surveys, transportation to the primary units is the chief cost item. The cost of adding a terminal or primary limb is much less than the cost of locating and driving to another block of trees. Thus, the most favorable variance situation would be the one in which the large variance component is the within-primary or terminal component. Not only is the cost of adding another terminal low, but the divisor of this component in the overall variance expression is a product of a number of elements of all levels. In addition, most primaries have only a small number of terminals and the terminal component is either eliminated by counting all terminals or greatly reduced by the within-primary finite correction factor. If trees within blocks are homogeneous, only a few trees per block would be needed. If stratification by age classes (or some other scheme) were effective, the block variance component would be reduced and fewer blocks would be needed.

Variance components were estimated by analysis of expanded estimates. Table 2 lists expanded means and variance components. Most data in table 2 are reasonable, but some unreasonable results were created by the analysis used. For example, the zero between-block component for Valencias, 0-3 years old, is a result of extremely high variation between trees within blocks, compared with the block average fruit counts. It is not reasonable to assume that no variation exists between blocks of trees of this type and age. Best interpretation of data for a particular category might be obtained by comparison of components with those indicated for other ages of the same type of citrus and other types of the same age. This approach was used for optimum allocation calculations.

Table 2.--Average fruit and variance components: Equal probability expansion, selected citrus blocks, Texas, 1969 growing season

Type and age class	Average fruit	Variance components 1/			
		Block	Tree	Primary	Terminal
Early oranges:					
0-3 years --	98	26,763	2,741	11,545	19,345
4-7 years --	430	22,224	36,267	117,908	136,149
8 years and over --	429	47,584	43,652	31,365	962,962
Grapefruit:					
0-3 years --	192	28,768	1,862	7,084	20,862
4-7 years --	229	25,130	0	5,510	163,917
8 years and over --	371	15,053	0	106,831	106,402
Valencia oranges:					
0-3 years --	182	0	4,272	37,702	22,466
4-7 years --	285	27,580	25,194	33,101	55,808

1/ Variance components indicated to be zero or negative by analysis of variance have been entered as zero.

TIME REQUIREMENTS

Average times from 1969 research provide the best available information on costs and may be used as a basis for estimating actual time requirements of an operational project.

Work on the 1969 project was performed mainly by college student trainees, SSO employees, and R&D personnel. Considering their background, it might be expected that these workers would carry out operations at a faster rate than persons who might be engaged for a large-scale survey. However, both survey periods were short (7-10 days), and workers were in a training phase, more or less, for most of the survey period. Consequently, average times should approach those of field personnel with more training and fieldwork experience.

Specific limb count operations which can be separated and analyzed include: (1) Time to measure and mark all primary limbs, (2) time required to map and mark terminal limbs on a primary, and (3) time required to count a terminal limb. Average time of each operation is available by age class and type of citrus.

Procedures used during July and September required that each primary limb be measured and marked with red flagging tape. Each terminal limb on designated primaries was measured and marked with blue flagging tape. All of these measurements were necessary to calculate and compare the possible estimators. All limbs were flagged to be sure that no primaries or terminals were missed and to identify terminals for selection of sizing limbs.

In a large-scale survey, all limbs would not have to be marked, although count or size limbs might require this type of identification. Also, if equal probability selection and expansion were used at the second stage, terminal limb CSA measurements would not be needed.

During the aerial photography phase of the regular research project, modified procedures were tried for selection operations. In this phase, a reliable estimate of fruit numbers on 20 large Valencia trees was needed. All primary limbs were measured and identification was made in chalk by drawing a line around each primary and marking it with a letter. The person doing the mapping moved in one direction around the trunk. Thus, when he returned to primary A, he could easily ascertain if all primaries were accounted for.

Terminals within primaries were determined by a terminal limb selection gauge instead of measurements. The gauge was 3- by 7-inch hard board with a 5/8-inch-deep opening cut in each end. At one end the opening corresponded to the diameter of a circle 0.6 square inch in area; at the other, to the diameter of a circle 1.2 square inches in area. In addition, a third notch was cut in one side corresponding to a circle 0.3 square inch in area.

A selected primary was followed out from the base. Each division qualifying as a terminal by the gauge was marked and numbered with chalk. A small limb (between 0.3 and 0.6 square inch) was combined with another of the same size or with a small terminal.

The terminal limb selection gauge proved very satisfactory. Generally, selection decisions were obvious. Most terminals were about 1.0 inches CSA. Occasionally, some limbs were more oval than circular--the gauge might fit across the small axis, but not the large axis. Most oval limbs tapered to a point which qualified on both axes before dividing.

Tables 3 and 4 summarize time required to identify and mark all first-stage primary limbs.

Table 3.--First-stage mapping: Minutes per tree, selected citrus blocks, Texas, 1969 growing season

Citrus type	Age class		
	0-3 years	4-7 years	8 years and over
Early oranges --	7.0	10.3	13.0
Grapefruit --	6.0	13.0	13.5
Valencia oranges --	15.3	11.6	1/ (10.5)

1/ Time reported is the average time for 20 trees from one block used for the work in November.

Table 4.--First-stage mapping: Minutes per primary, selected citrus blocks, Texas, 1969 growing season

Citrus type	Age class		
	0-3 years	4-7 years	8 years and over
Early oranges --	0.93	1.11	1.23
Grapefruit --	1.38	1.49	1.65
Valencia oranges --	2.16	1.64	1/ (.72)

1/ Time reported is the average time for 20 trees from one block used for the work in November.

Comparison with similar age trees of other types of citrus indicates about 2.5 to 3.0 minutes per tree can be saved by chalking primaries instead of using flagging tape. Mapping time was 12 minutes or less for 17 of the 20 large Valencia trees.

Average time for Valencias, 0-3 years, was very high. This might be attributed to inexperience, since one of the three blocks of this age was the first block worked during the survey period.

Greatest gain from new procedures would be in time saved identifying terminal limbs on primary limbs. Ordinarily, about 7.5 minutes are required to map Valencia primaries, 8 years and over, into terminals. Use of the terminal limb selection device and chalk reduced the average time per primary to 1.7 minutes. Tables 5 and 6 summarize average time of second-stage mapping by age and type of citrus.

Table 5.--Second-stage mapping: Minutes per primary, selected citrus blocks, Texas, 1969 growing season

Citrus type	Age class		
	0-3 years	4-7 years	8 years and over
Early oranges --	6.9	4.6	5.2
Grapefruit --	3.2	8.4	8.6
Valencia oranges --	4.1	7.2	1/(1.7)

1/ Limb selection gauge used to determine terminal limbs instead of measuring cross-sectional area of each terminal for 20 trees from one block used for the work in November.

Table 6.--Second-stage mapping: Minutes per terminal, selected citrus blocks, Texas, 1969 growing season

Citrus type	Age class		
	0-3 years	4-7 years	8 years and over
Early oranges --	2.20	0.97	1.02
Grapefruit --	.90	1.52	1.02
Valencia oranges --	1.32	1.80	1/(.34)

1/ Limb selection gauge used to determine terminal limbs instead of measuring cross-sectional area of each terminal for 20 trees from one block used for the work in November.

The third major operation in a limb count survey is actual counting time of limbs. Tables 7 and 8 summarize average counting times per terminal limb and per fruit in the 1969 research and should be studied together. Average counting time per terminal of Valencias, 8 years and over, in table 7 is less than might be expected based on other Valencia trees. In addition, table 8 shows that average counting time per fruit was the lowest for Valencias, 8 years and over. Thus, more fruit were counted in less time. No new procedures were used in the actual counting procedure. The fact that earlier (mapping and selection) operations did not take as much time probably left workers fresher for limb counting.

Table 7.--Fruit counting: Minutes per terminal limb, selected citrus blocks, Texas, 1969 growing season

Citrus type	Age class		
	0-3 years	4-7 years	8 years and over
Early oranges --	2.8	2.4	2.4
Grapefruit --	1.6	1.9	2.0
Valencia oranges --	2.1	2.3	1/(1.4)

1/ Time reported is for the 20 trees from one block used for the work in November.

Table 8.--Fruit counting: Minutes per fruit, selected citrus blocks, Texas, 1969 growing season

Citrus type	Age class		
	0-3 years	4-7 years	8 years and over
Early oranges --	0.183	0.234	0.307
Grapefruit --	.166	.332	.346
Valencia oranges --	.283	.246	1/(.132)

1/ Time reported is for the 20 trees from one block used for the work in November.

On the average, a two-man team will require about 25-26 minutes to select limbs and count fruit on large trees--10.5 minutes for first-stage mapping, 3.5 minutes to map two primaries, 9 minutes to count six terminal limbs, and 2-3 minutes for between-operation gaps. Smaller trees might be worked even faster. Times would be slightly greater if primary and terminal limbs were marked for followup studies.

OPTIMUM ALLOCATION

In the 1969-70 research project, four trees per block were allocated for limb counts and photography to ensure good estimates of the within-block and within-tree variance components. At least two observations were made at all levels of sampling (groves, trees, primary limbs, and terminal limbs) so corresponding variances could be calculated.

The data collected on variance components and variable costs (times) allowed calculations of optimum number of subsamples to be determined for all levels of sampling. The variance model assumed for estimating fruit per tree ($\bar{Y}$) from limb counting was

$$\sigma_{\bar{Y}}^2 = \frac{\sigma_b^2}{n} + \frac{\sigma_t^2}{na} + \frac{\sigma_p^2}{nab} + \frac{\sigma_w^2}{nabc} ,$$

where

σ_b^2 is the between block or grove variance component,

σ_t^2 is the between tree variance component,

σ_p^2 is the between primary limb variance component, and

σ_w^2 is the between terminal limb variance component.

The divisors n, a, b, and c refer to the number of blocks, trees per block, primary limbs per tree, and terminal limbs per primary, respectively. The variable cost function assumed for this model is

$$\text{Total variable cost} = n\,C_1 + na\,C_2 + nab\,C_3 + nabc\,C_4$$

where

C_1 = Cost of adding each additional block of fruit. Cost is mainly travel time and mileage, but includes extra preparation time and cost, extra supplies and time for field crew to organize forms and equipment at each block.

C_2 = Cost of adding each additional tree. This cost includes time to locate designated tree plus time to identify and measure all primary limbs; that is, the cost of getting ready for the next stage of sampling.

C_3 = Cost of adding each additional primary limb. This is the time required to mark the selected primary limb and identify all terminal limbs on it. The time required to count path fruit is considered a primary limb cost.

C_4 = Cost of adding each additional terminal limb. This is the time required to mark the selected terminal limb and to count the fruit on the limb.

Given the variance and cost assumptions above, optimum values can be found for a, b, and c by the formulas:

$$\text{Optimum } a = \sqrt{\frac{C_1}{C_2} \cdot \frac{\sigma_t^2}{\sigma_b^2}}, \qquad \text{Optimum } b = \sqrt{\frac{C_2}{C_3} \cdot \frac{\sigma_p^2}{\sigma_t^2}}$$

$$\text{Optimum } c = \sqrt{\frac{C_3}{C_4} \cdot \frac{\sigma_w^2}{\sigma_p^2}}$$

The optimum value for n, the number of blocks, can then be found by solving the variance formula for a desired variance or the cost function for a fixed total variable cost.

Separate analyses were performed for each citrus type and age combination. Negative or zero variance components were adjusted based on the relationship of other types and ages. In addition, variance components were calculated for the combination of the 4-7 years and 8-years-and-over age categories.

In estimating costs, it was assumed that a previous visit had been made to each block so the location of blocks and trees within blocks was already known. This assumption is realistic because the sampling frame would need to be improved during the first few years of an operational project by making visits to verify tree numbers, ages, and types. Selection of sample blocks from this frame evaluation sample would eliminate the time cost of tree selection during the fruit estimation survey itself.

A value of 60 minutes was used for the between-block cost. In addition to driving time between blocks, this cost includes allowances for preparation costs, block-to-home travel, and time from car to first tree. Second-stage cost was estimated essentially as listed in the time requirements section. It was assumed that all primary limb CSA's would be measured and recorded, but marking would be done with chalk at this time.

Costs for the third stage (primary limbs) were estimated at 3.0 or 3.5 minutes for equal probability expansions, assuming no measurement of terminal limbs. In all cases, it was assumed that each selected primary limb would be marked with flagging tape or another more permanent device before terminal limbs were identified. It was also assumed that path fruit would be counted at this time.

Counting time (C_4) was estimated at 4 minutes for each terminal limb. This includes marking the limb with flagging tape or some other device and allows time to recheck any doubtful counts. An experienced crew should take about 3 minutes for most counts, but some limbs will be rather inaccessible and will require more time.

Table 9 lists the optimization results for equal probability expansions. Indicated optimum number of trees per block ranges from 1.28 to 2.41. Therefore, the sample would probably consist of two trees per block for each type and age category. If there were another objective besides yield, such as determining fruit quality or increasing efficiency through photography, more than two trees might be required for the sample.

In general, a sample assignment of two primaries per tree and two terminals per primary will be close to the optimum. Some of the results which indicate a larger sample size may be due to sampling variation inherent in data for the 1969 season.

Table 9.--Optimization results: Equal probability expansions, selected citrus blocks, Texas, 1969 growing season

Citrus type and age	Optimum allocation values		
	Trees per block	Primaries per tree	Terminals per primary
Early oranges:			
0-3 years	1.28	2.37	1.12
4-7 years	2.07	3.67	1.00
4 years and older	2.16	2.87	1.82
8 years and older	2.28	2.03	2.96
Grapefruit:			
0-3 years	1.29	1.67	1.50
4-7 years	1.52	1.22	4.89
4 years and older	1.46	4.41	1.23
8 years and older	1.46	5.34	.94
Valencia oranges:			
0-3 years	1.66	2.42	.93
4-7 years	2.41	1.77	1.22
4 years and older	2.33	2.02	1.51
8 years and older	2.30	2.15	1.62

The number of blocks needed for specific levels of variation can be estimated by solving the equation

$$\sigma_{\bar{Y}}^2 = (1-f_1)\frac{\sigma_b^2}{n} + (1-f_2)\frac{\sigma_t^2}{na} + (1-f_3)\frac{\sigma_p^2}{nab} + (1-f_4)\frac{\sigma_w^2}{nabc}$$

for n, where f_1, f_2, f_3, and f_4 are the finite correction factors at the respective levels of sampling. The finite correction factor, f_i, may be defined as number sampled at level i divided by the average total number in the population at level i. The finite correction factors are of no concern (less than 1 percent) at the block and tree level, but are important at the primary and terminal sampling level. Tables 10 and 11 list the average number of primaries per tree and terminals per primary from the 1969-70 survey.

Table 10.--Average number of primary limbs per tree, selected citrus blocks, Texas, 1969 growing season

Citrus type	Age class		
	0-3 years	4-7 years	8 years and over
Early oranges --	4.615	8.286	11.129
Grapefruit --	5.909	8.857	10.941
Valencia oranges --	7.083	6.943	------

Table 11.--Average number of terminal limbs per primary, selected citrus blocks, Texas, 1969 growing season

Citrus type	Age class		
	0-3 years	4-7 years	8 years and over
Early oranges --	2.900	4.559	4.974
Grapefruit --	2.957	4.957	6.296
Valencia oranges --	3.167	4.053	-----

A common procedure in calculating number of blocks to sample is to set the degree of confidence required of the survey. If the desired error is set at 10 percent of the mean with a 67-percent confidence interval, the equation is: $\sigma_{\bar{Y}}^2 = (.1\,\bar{Y})^2$.

The optimization of n_i, the number of blocks within an age class, can be approached in two different ways. If estimates are desired by age class as well as by the total of all classes, a confidence level must be set for each class and for the total population. This is termed a multiple-purpose optimum allocation. If the goal is simply an estimate of the total fruit of each type, a confidence level is set only for the total and it becomes a standard single-purpose optimum allocation. Estimate of average fruit per tree for a particular type is given by

$$Y_i = W_1 \bar{Y}_1 + W_2 \bar{Y}_2 + W_3 \bar{Y}_3,$$

with $W_h = N_h/N$, with N_h the number of trees in the h^{th} age class and N the total number of trees of the type.

$$\sum_{h=1}^{3} W_h = 1$$

Variance of this estimate is

$$\sigma_{\bar{Y}}^2 = W_1^2 \sigma_{\bar{Y}_1}^2 + W_2^2 \sigma_{\bar{Y}_2}^2 + W_3^2 \sigma_{\bar{Y}_3}^2 .$$

Since the original sampling frame was 3 years old, a 1969 listing of tree numbers was used for weighing. In this listing, age categories 4-7 years and 8 years and over were combined. Estimates of variance components for this combination were used to calculate sample size requirements. Table 12 summarizes the calculations for a 10-percent coefficient of variation for each type of fruit.

The steps necessary to calculate total n and the values of n_1 and n_2 are outlined in appendix B. All calculations are shown for the early orange data.

Table 12.--Calculations of n_h for each type of citrus, selected citrus blocks, Texas, 1969 growing season

Type	Average fruit per tree 1/	Variance required 2/	Total n required	Allocation		Coefficient of variation	
				n_1	n_2	Class 1	Class 2
						Percent	Percent
Early oranges --	272.13	740.547	85	21	64	39.22	8.94
Grapefruit --	268.93	723.233	79	28	51	18.23	11.67
Valencia oranges --	301.55	909.324	65	21	44	21.58	11.22

1/ $\bar{Y} = W_1 \bar{Y}_1 + W_2 Y_2$, where 1 refers to 0-3 age class and 2 refers to 4 years and older.

2/ $\sigma_{\bar{Y}}^2 = (.1 \bar{Y})^2$.

Most analyses of 1969 fruit counts employed the equal probability expansions discussed in the expansion of terminal counts section on page 7 with results shown in table 2. Two other two-stage estimators were calculated for comparison with the equal probability expansions. These estimators make use of relative limb size at the primary sampling level or at both sampling levels.

One possible estimator is referred to as a double ratio estimator. For each terminal, the ratio of fruit count to terminal CSA is computed and multiplied by the CSA of all terminals on the primary. This factor is then multiplied by the total CSA of all primary limbs divided by the CSA of the selected primary limb. In equation form, the estimate of total fruit per tree from the ith terminal of the jth primary within a tree is

$$(\hat{y}_{ij})_2 = \frac{\sum_{j=1}^{m} A_j}{A_j} \cdot \frac{\sum_{i=1}^{n_j} a_{ij}}{a_{ij}} \cdot X_{ij}$$

where

$\hat{y}_{ij}$ is estimated fruit per tree,

x_{ij} is number of fruit counted on the terminal,

m is the number of primary limbs,

n_j is the number of terminals on the jth primary,

A_j is CSA of the jth primary, and

a_{ij} is the CSA of the ith terminal on the jth primary.

This estimator is referred to as a double ratio estimator since at each level the ratio of the total area at that level to the area of the selected limb is used. It is not an unbiased estimator of total fruit per tree unless the relationship between fruit counted and CSA passes through the origin.

Another possible estimator might be labeled a combination ratio/cluster expansion. A ratio estimator of CSA of all primary limbs divided by the selected primary CSA is used at the first stage, but the second-stage expansion is the number of terminal limbs on the selected primary. The estimate of total fruit per tree from the ith terminal of the jth primary within a tree is

$$(\hat{y}_{ij})_3 = \frac{\sum_{j=1}^{m} A_j}{A_j} \cdot n_j \cdot x_{ij}$$

where

$\hat{y}_{ij}$ is the estimate of fruit per tree,

x_{ij} is the number of fruit counted on ith, jth terminal,

n_j is the number of terminals on the jth primary,

m is the number of primary limbs, and

A_j is the CSA of the jth primary.

This estimator is again a ratio estimator and not necessarily unbiased. Each primary is assumed to be a cluster of terminal units of equal size. This is a reasonable assumption since terminal limb size was limited to 0.6 to 1.2 square inches CSA. With such a narrow range, correlation of terminal limb size and fruit counts would not be very high.

Like the equal probability expansions, fruit counts were expanded for all terminal limbs (with proper adjustment for path fruit) and analyzed for components of variation. These results are presented in tables 13 and 14.

Table 13.--Average fruit variance components: Double ratio estimator, selected citrus blocks, Texas, 1969 growing season

Citrus type and age class	Trees	Average fruit	Variance components 1/			
			Block	Tree	Primary	Terminal
	Number	Number				
Early oranges:						
0-3 years	16	96	26,075	1,850	12,874	15,584
4-7 years	36	384	33,249	19,476	54,361	68,068
8 years and over	36	393	45,503	40,640	0	169,335
Grapefruit:						
0-3 years	12	191	29,729	1,716	3,658	11,250
4-7 years	36	198	15,300	3,811	4,962	60,675
8 years and over	36	338	943	14,553	40,497	67,717
Valencia oranges:						
0-3 years	12	183	198	10,949	7,846	15,137
4-7 years	36	274	32,709	30,912	6,782	42,609

1/ Variance components indicated to be zero or negative by analysis of variance have been entered as zero.

Table 14.--Average fruit variance components: Ratio/cluster estimate, selected citrus blocks, Texas, 1969 growing season

Citrus type and age class	Trees	Average fruit	Variance components 1/			
			Block	Tree	Primary	Terminal
	Number	Number				
Early oranges:						
0-3 years	16	94	25,951	0	14,786	16,550
4-7 years	36	395	36,794	24,362	47,753	98,260
8 years and over	36	394	29,673	49,296	0	234,834
Grapefruit:						
0-3 years	12	194	31,063	2,550	0	18,660
4-7 years	36	201	14,424	3,165	0	80,779
8 years and over	36	349	919	15,918	31,783	92,432
Valencia oranges:						
0-3 years	12	174	0	8,700	9,322	15,080
4-7 years	36	276	31,804	32,633	0	60,398

1/ Variance components indicated to be zero or negative by analysis of variance have been entered as zero.

Data in tables 13 and 14 are comparable to the equal probability results in table 2. As in table 2, some variance components are not reasonable due to sampling variation. Estimates of actual variance components can be made by comparing results for various types and ages.

To compare the results of the three estimators, the coefficient of variation (standard deviation of the estimate of the mean divided by the mean) was calculated for the various fruit types. Sample sizes in the 1969 research work were used to estimate the variances and standard deviations. The formula used for calculating variance was

$$S_{\bar{y}}^2 = \frac{S_b^2}{n} + \frac{S_t^2}{na} + (1-\bar{f}_3 \; \frac{S_p^2}{nab} + (1-\bar{f}_4 \; \frac{S_w^2}{nabc}$$

where

$S_{\bar{y}}^2$ is calculated variance of the estimated mean,

S_b^2 is the block variance,

S_t^2 is the tree variance,

S_p^2 is the primary limb variance,

S_w^2 is the terminal limb variance,

$(1-f_3)$ is the finite correction factor at the primary limb level,

$(1-f_4)$ is the finite correction factor at the terminal level,

n is the number of blocks sampled,

a is the number of trees per block,

b is the number of primary limbs sampled per tree, and

c is the number of terminal limbs per primary limb.

Table 15 presents the results of the calculations for the three estimates.

A slight reduction in variance for most age classes through the use of either the double ratio or ratio/cluster estimator is shown in table 15. These gains from use of limb size in estimation are not as great as might be expected. The reductions in variation from use of limb size data came mainly at the primary and terminal level. The variance due to differences between blocks and between trees within blocks is not affected.

The effects of variance components at the primary and terminal limb level are reduced by large divisors and by finite correction factors. Thus, a considerable reduction in variance (40-50 percent) at the primary and terminal levels might reduce the coefficient of variation less than 1 percent.

Use of the double ratio estimator might actually be less efficient than equal probability estimation when costs are considered. Equal probability or ratio/cluster estimation required identification of terminal limbs. Measurement of each terminal limb, as required by double ratio estimation, might take 4 to 5 extra minutes per primary limb.

Table 15.--Coefficients of variation of estimated average fruit count for three possible estimators, selected citrus blocks, Texas, 1969 growing season

Citrus type and age class	Equal probability	Double ratio	Ratio/cluster
	Percent	Percent	Percent
Early oranges:	85.73	85.80	86.02
0-3 years	13.64	12.65	12.89
4 years and over			
Grapefruit:			
0-3 years	53.83	53.61	54.03
4 years and over	16.00	15.02	15.18
Valencia oranges:			
0-3 years	53.96	52.92	53.00
4 years and over	14.55	14.21	14.81

GROUND PHOTOGRAPHY

Nearly every tree on which fruit counts were made was photographed in the 1969-70 project. Blocks which would have required a special visit or which had little fruit were not photographed.

Most pictures were made with high-speed Ektachrome film exposed in a Miranda or Minolta single-lens reflex camera with a light meter behind the lens. Some Kodachrome II film was also used.

Photos were taken of only one side of each tree. Previous research showed no statistical differences between counts from photography of different sides of trees.

An aluminum frame (upright and crossbar) was used to divide large trees into upper and lower quadrants for photography. Only one or two photos were taken of smaller trees, when the entire side or one-half of the side could be photographed in one photo. All photos were made at a distance of 15-20 feet from the tree.

Photography instructions were nearly the same as those outlined in appendix B of the 1968 summary report. 2/ One modification which helped improve the quality of photos was to adjust the light meter indication when shooting upper portions of trees. Underexposed photos were sometimes obtained when a change of two or more f-stops was made, as indicated by the light meter reading. By changing only one f-stop, a better exposure was achieved. Light

---

2/ Richard D. Allen and Donald H. Von Steen. Use of Photography and Other Objective Yield Procedures for Citrus Fruit, 1968 Texas Research, U.S. Dept. Agr., Statis. Rptg. Serv., June 1969.

meter readings for lower quarters gave generally good results without modification.

Most photos of early oranges and grapefruit were taken in July. Valencia photos and some shots missed in July were taken in September. September photography was generally better for fruit counting than that in July, apparently due to differences in sun angle and light intensity.

Almost all slides for the blocks selected in 1969 were interpreted by the same counter for the following reasons:

(1) To have counts available as soon as possible;
(2) To have counts as comparable as possible from slide to slide;
(3) To estimate photo counting times for an individual familiar with citrus trees.

After obtaining fruit counts, correlation analyses were performed to compare estimated fruit per tree and fruit counted from photography. To compensate for differences in slides per tree, all photo counts were converted to total per side of tree before analyses were performed. These correlation results are listed in table 16 with equivalent percentages of estimated fruit counted.

Table 16.--Comparison of photo counts, selected citrus blocks, Texas, 1968 and 1969 growing seasons 1/

Type	Percentage counted		Correlation (5)	
	1968	1969	1968	1969
	Percent	Percent		
Early oranges	30.9	25.4	0.798**	0.684**
Grapefruit	32.4	19.3	.915**	.617**
Valencia oranges	31.0	29.1	.546	.715**
All oranges	29.9	25.2	.640**	.680**
Oranges, trunk size less than 80 inches CSA	36.0	26.1	.794**	.708**
Oranges, trunk size greater than 80 inches CSA	23.7	22.0	.517	.456

1/ 1968 data are for 8 research blocks and are based on actual tree counts. 1969 data are for random selection of blocks using estimated tree totals. The 1969 correlations have not been corrected for the sampling errors of estimated tree totals.
** Indicates that correlation is greater than zero with p = 0.99.

Correlations of counts from photography with estimated total fruit per tree are very encouraging. Although it is reasonable to expect a much lower correlation of photo counts with estimated counts than with actual counts, correlations from 1969 data compare favorably with 1968 results.

To benefit from using an auxiliary variable such as photo counts of fruit in a regression or unbiased ratio expansions, correlation (r) should be greater than $1/2\ (S_{\bar{x}})\ /\ \bar{X} \div (S_{\bar{y}})\ /\ \bar{Y}$. 3/ For grapefruit, this calculation gives a value of 0.500, less than the sample correlation. For orange trunks, exceeding 80 inches CSA, calculated value of 0.632 is greater than the sample value. Hence, photo counts might not be desirable in this instance.

If a double sampling approach within blocks is used to combine photo count and limb count, one procedure might be to make the limb counts for two trees, as indicated by optimum allocation, but photograph a larger number of trees, including the two with limb counts. The estimating model for the ith block of fruit would be

$$\hat{Y}_i = \bar{Y}_i + b\ (\bar{X}_{ia'} - \bar{X}_{i2})$$

where

$\hat{Y}_i$ is estimated number of fruit per tree in the i^{th} block,

$\bar{Y}_i$ is the direct expansion estimate from the limb counts on two trees within the block,

b is the slope of the regression line of photo counts on estimated actual counts for same trees,

$\bar{X}_{ia'}$ is average number of fruit counted from photos for the a′ trees photographed within the block, and

$\bar{X}_{i2}$ is average number of fruit counted from photos for the two trees with both limb and photo counts.

Variance function within blocks is as follows:

$$\text{Variance} = \frac{\sigma_t^2\ (r^2)}{a'} + \frac{\sigma_t^2\ (1-r^2)}{2} + \frac{\sigma_p^2}{4} + \frac{\sigma_w^2}{8}$$

where the variance components are those defined throughout this report and sampling rates within trees are two primaries and two terminals per primary. Given a cost function of TC = a′ C′ + n C for the photo counting and limb counting portions, respectively, the optimum value for the ratio of trees to

---

3/ Des Raj. Sampling Theory. McGraw-Hill Book Co., New York, 1968, p. 92.

photograph is

$$\frac{a'}{n} = \sqrt{\frac{C\ r^2}{C'\ (1-r^2)}}$$

Time cost per tree for photo counts within blocks having limb counts would be approximately half the cost of limb counting. Allowance is made for film and processing costs, but it is assumed that photo counting is done by lower paid workers. If the ratio of C to C′ is 2 and n = 2, the result is

$$a' = 2\sqrt{\frac{2\quad r^2}{(1-r^2)}} \quad \doteq 3\sqrt{\frac{r^2}{(1-r^2)}}$$

Since most correlations from the photo counting are about 0.7, the ratio of r^2 to $(1-r^2)$ is approximately 1 and optimum value of a′ would round to three trees per block.

Correlation figures in table 16 were calculated by using all x and y values. Within-block correlation figures should be higher and double sampling would be more efficient. (Within-block correlation (r) was higher than this population figure in nine out of 11 blocks of Valencia oranges.) However, within-block regression coefficients cannot be calculated if limb counts are made on only two trees per block.

To test the suitability of using a single regression coefficient, the individual block regressions can be compared. A summary of this comparison for Valencia photo counts is shown in table 17. 4/

The hypothesis that only one slope is needed is tested by an F-test of the additional deviations from a single regression line divided by the sum of deviations from individual block regressions. In this case, F = 25,381 ÷ 23,236 = 1.092. This value is not significant and the hypothesis of common slope is accepted. Next is a test of whether a common intercept could be used. The test of common intercept is the difference between adjusted means divided by the deviations from pooled regression. In this case, F = 32,225 ÷ 23,910 = 1.348, which is not significant. Thus, it is concluded that a single regression line of the form $Y_{ij} = a + b\ X_{ij}$ could be used for all Valencia photo counts.

Instead of photographing extra trees within a block, photo counts only might be made of some blocks. This would allow more effective utilization of time and personnel. More experienced workers could take the photos and make the necessary limb counts; less experienced, lower paid workers might be employed for the photo counting. Photo counting, unlike limb counting, is not limited to daylight hours on good days.

4/ Procedure is taken from George W. Snedecor and William G. Cochran. Statistical Methods. Iowa St. Univ. Press, Ames, Sixth ed., 1967, pp. 432-436.

Table 17.--Comparison of photo count regression lines, Valencia oranges, selected citrus blocks, Texas, 1969 growing season

Source of variation	Deviations from regression		
	Degrees of freedom	Sums of squares	Mean squares
Sum of individual block deviations	24	557,658	23,236
Deviations from pooled regression	35	836,853	23,910
Difference between slopes	11	279,195	25,381
Deviations from between plus within	46	1,191,332	25,899
Difference between adj. means	11	354,479	32,225

Tables 18 and 19 list average counting times and average counts per slide from 1969 slides. Since some slides showed half or all of one side of a tree instead of a quarter, types of slides are listed separately.

The breakdown of early oranges and grapefruit refers to different counting assignments. Blocks containing mostly large trees were generally photographed by quarters; blocks with a mixture of the three types of photography were grouped together in other assignments.

Average counting time per slide in 1969 was less than in 1968 when clerical employees did the counting. Average 1968 times per slide ranged from 5 to 9 minutes, with times of 5 to 7 minutes for slides with averages of 25 or fewer fruit. Two sets of 1969 photos were counted by the clerical staff. Results were an average time of 6.4 minutes and 14.1 fruit per slide for early oranges, and 4.4 minutes and 9.3 fruit per slide for grapefruit.

Much of the difference in the two levels of counts can be credited to experience with citrus and counter motivation. A person who has worked with citrus in the field can identify fruit more readily and make identification decisions faster. To avoid sacrificing accuracy for speed, all counts in 1969 were checked by scanning across the slides in the direction opposite to that taken in the original counting.

It is not suggested that a group of counters could be trained to count as quickly as the times shown in table 18. People with knowledge of the citrus

Table 18.--Average fruit counting time per slide and number of slides per portion of tree counted, selected citrus blocks, Texas, 1969 growing season

Citrus type	Portion of tree: Entire side: Slides	Entire side: Average counting time	Half of side: Slides	Half of side: Average counting time	Quarter of side: Slides	Quarter of side: Average counting time
	Number	Minutes	Number	Minutes	Number	Minutes
Early oranges:						
Small trees	13	1.692	32	2.344	8	1.375
Large trees	--		2	3.500	30	2.300
Grapefruit:						
Small trees	6	2.167	11	1.636	23	1.652
Large trees	--		9	2.222	81	1.741
Valencia oranges	8	2.875	50	2.860	60	2.667

Table 19.--Average fruit counted per slide and number of slides per portion of tree counted, selected citrus blocks, Texas, 1969 growing season

Citrus type	Portion of tree: Entire side: Slides	Entire side: Average counting time	Half of side: Slides	Half of side: Average counting time	Quarter of side: Slides	Quarter of side: Average counting time
	Number					
Early oranges:						
Small trees	13	4.692	32	14.281	8	5.375
Large trees	--		2	27.500	30	16.100
Grapefruit:						
Small trees	6	7.667	11	6.091	23	7.087
Large trees	--		9	13.000	81	7.605
Valencia oranges	8	10.625	50	14.600	60	15.567

crop hired especially for the counting should average 3 or 4 minutes per slide instead of 5 to 9 minutes, however.

COMPARISON OF ESTIMATING SYSTEMS

To compare possible estimating systems (using auxiliary information along with fruit counting), calculations were made from 1968 research data. Actual counts of total fruit per tree were available for a total of 32 trees. Fruit counts and limb measurements were recorded for each limb. Estimators considered (expressed as estimate for ith tree) included:

Equal probability: $\hat{Y}_1$ = total number of limbs $\frac{\text{Total fruit on sizing limbs}}{\text{Total number of sizing limbs}}$

Proportional to size: $\hat{Y}_2$ = Average of: $\frac{\text{Fruit on sizing limb}}{\text{CSA of sizing limb}}$

Trunk size regression: $\hat{Y}_3 = \hat{Y}_1 + b_1 (X_i - \bar{X})$

Primary size regression: $Y_4 = Y_1 + b_2 (X_i - \bar{X})$

Photo count ratio: $\hat{Y}_5 = (\text{Photo count})_i \cdot \frac{1}{R}$

Photo count regression: $\hat{Y}_6 = \hat{\bar{Y}}_1 + b_3 (X_i - \bar{X})$

Composite photo/limb count: $\hat{Y}_7 = (\text{Photo count})_i + (1-R)\ \hat{Y}_i$

Regression coefficients b_1, b_2, and b_3 are calculated from regression analyses of individual tree estimates from estimator $\hat{Y}_1$ and x variables; trunk CSA, total primary CSA, and photo count for the ith tree, respectively, The ratio R in estimators 5 and 7 is total photo count for the set of trees divided by total fruit from $\hat{Y}_1$.

Estimators 3, 4, 5, and 6 are the type of estimators used in double sampling where the parameters ($\bar{Y}$'s, b's, and R) are calculated from a small set of data and estimates are made for a larger set of data for which the x variable only is measured. Since all data were used in calculation of parameters, the estimates $\hat{Y}_3$ through $\hat{Y}_7$ are not independent of $\hat{Y}_1$. If the estimates $\hat{Y}_3$ through $\hat{Y}_7$ had been calculated based on $\hat{Y}_2$, the errors of the estimates probably would have been smaller for oranges. (See table 20.) The objective in testing is to determine whether additional information from the sample trees will give a "better" estimate than fruit count along. "Better" is defined as having a smaller squared difference from the true value for the tree.

Results from these calculations are shown in table 20. Calculations were made for 32 citrus trees--16 early orange, eight Valencia, and eight grapefruit.

Table 20.--Comparison of various estimators, selected citrus blocks, Texas, 1969 growing season

Estimator	Variance of estimated values from actual counts		
	Early oranges	Valencia oranges	Grapefruit
Equal probability	27,461	81,223	5,215
Proportional to size	18,575	27,082	7,953
Trunk size regression	68,146	126,465	15,838
Primary size regression	49,653	50,158	16,190
Photo count ratio	25,810	131,311	3,588
Photo count regression	25,863	142,036	4,463
Composite photo limb count	22,211	67,694	3,898

The figures in table 20 were calculated by the formula:

$$\text{Variance (jth estimator)} = \frac{\sum_{i=1}^{n} (m_i - 1)(\hat{X}_{ij} - X_i)^2}{\sum_{i=1}^{n} (m_i - 1)}$$

where

X_i is the actual fruit count for the ith tree,

$\hat{X}_{ij}$ is the estimate of the jth estimator for the ith tree,

m_i is the number of sizing limbs represented in $\hat{Y}_{1i}$ and $\hat{Y}_{2i}$, and

n is the number of trees in the data set.

The proportional-to-limb-size results were closer to actual counts than equal probability results for two of the three types. Average of the expanded data was closer to the true average fruit count than the average of the equal probability expansions of these two types.

Use of trunk size in a regression estimator did not improve the closeness of the estimator for any of the three types of fruit. Regression estimation using the total CSA of all primary limbs did not reduce the variation in two of the three types, but it gave Valencia oranges the smallest variation from actual counts.

The composite photo count estimator, which uses actual photo count and allows for portion uncounted, was the most consistent of the estimators. It always gave a lower variation than the equal probability estimator and always had the lowest or second lowest variation from actual count of any of the six estimators (excluding probability proportional to size).

Photo count regression and ratio estimators both performed better than equal probability expansion for grapefruit and early oranges, but not for Valencia oranges. The low photo count percentage of a block of large Valencia trees may have been a factor in the poor showing of the photo count ratio and regression estimators.

Historic parameters may also form the basis for an estimation system. That is, model parameters based on relationships derived in previous years may be used with current photo counts or limb counts to estimate fruit per tree. This method assumes that basic relationships are fairly stable and that accumulation of parameter values over time will lead to smaller errors in the estimators and will reduce the cost of collecting information.

Data from the 1969 random sample of blocks were used to calculate these historic parameters. The results were then used to predict tree totals of trees sampled in 1968.

Four estimators applying historic parameters--trunk size regression, total primary CSA regression, photo count ratio, and photo count regression--are compared with estimators using only current (1968) data in table 21.

Table 21.--Comparison of current and historic estimators, selected citrus blocks, Texas, 1968 growing season

Estimator	Variance of estimated values from actual counts		
	Early oranges	Valencia oranges	Grapefruit
Trunk size regression			
Current	68,146	126,465	15,838
Historic	222,494	275,547	17,441
Primary size regression:			
Current	49,653	50,158	16,190
Historic	209,392	59,911	19,121
Photo count ratio			
Current	25,810	131,311	3,588
Historic	104,445	335,578	7,546
Photo count regression:			
Current	25,863	142,036	4,463
Historic	23,169	158,898	13,404

Historic parameters for early season and Valencia oranges resulted in much larger errors for all but the photo count regressions. Average fruit counts from the 1969 random sample of early orange and Valencia orange blocks were considerably lower than for the sample blocks used in 1968. The parameters from the 1969 grapefruit study showed marked similarity when applied to the 1968 grapefruit data.

From the foregoing comparisons, it is apparent that estimators based on current season data are preferable to historic parameters, unless larger errors can be accepted. Equal probability expansion of terminal fruit counts is also preferred because of lower cost and simplicity of sampling attributes.

Fruit counts from photography can be used in the estimation system in a double sampling scheme in which some blocks have both photo and limb counts. The relationships from these blocks can then be applied to additional blocks having only photo counts.

Measurements of trunk CSA are apparently of limited value in the estimation system. However, if these measurements can be obtained in great quantity and at little cost (during the updating of the tree numbers frame, for example), they might prove useful.

If several estimators of fruit per tree are used, they should be combined by the best linear estimator technique. 5/ This procedure will give the smallest variance in estimating fruit per tree for each type of citrus.

FRUIT SIZE STUDY

Main indications of fruit size development during 1969-70 came from the eight research blocks used in 1968-69. Each of these blocks was visited monthly (except February) until harvest. Fruit were tagged in the random sample of research blocks to provide estimates of components of variance. These new blocks were revisted whenever possible, but less frequently than the eight blocks.

Monthly fruit sizes for the 1969 growing season are shown in tables 22, 23, and 24, with comparisons to the previous season. Four-digit numbers following the special research summary data refer to individual averages from the new research blocks.

Two types of comparisons can be made from tables 22, 23, and 24: (1) comparisons of the same blocks between years and, (2) comparisons of 1969-70 averages of randomly selected blocks with nonrandom blocks. Plotting the averages on graphs aided in analysis of these comparisons.

After September 1, 1969, growth rate of early oranges (average of four blocks) was approximately the same as a year earlier during the same period. A big increase in size during August may have been a response to irrigation or weather conditions. Much lower fruit counts of Marrs block II and pineapple oranges in 1969 contributed to increased size of fruit in these blocks.

5/ Sampling Theory, pp. 16-17.

Average fruit size increased faster in the special research blocks than in the nonrandom early orange blocks. Generally, the special research blocks contained fewer fruit per tree.

Table 22.--Early oranges: Average diameter size, selected citrus blocks, Texas, 1968 and 1969 growing seasons 1/

Block	Year	Aug. 1	Sept. 1	Oct. 1	Nov. 1	Dec. 1
		Inches				
Marrs I 2/	1968	1.96	2.20	2.40	2.63	
	1969	1.86	2.25	2.45	2.58	2.65
Marrs II	1968	2.03	2.20	2.39	2.50	
	1969	2.12	2.38	2.55	2.74	2.89
Pineapple	1968	2.18	2.33	2.42	2.56	
	1969	2.15	2.38	2.56	2.64	2.69
Jaffa	1968	1.96	2.11	2.26	2.41	2.52
	1969	2.04	2.21	2.40	2.52	2.52
Average, 4 blocks	1968	2.03	2.21	2.37	2.52	
Average, 4 blocks	1969	2.04	2.30	2.49	2.62	2.69
Special research 3/	1969	2.09		2.58	2.74	
1101	1969	1.84		2.39	2.57	
1205	1969	1.77		2.34		
1208	1969	1.91		2.43		
1210	1969	2.29		2.77		
1303	1969	1.84		2.26	2.42	
1306	1969	1.92		2.35	2.50	
1310	1969	1.95		2.43		

1/ Data for the four blocks and the special research averages have been converted to equivalent first-of-month sizes. All data are average diameter in inches.

2/ Average sizes for November and December 1, 1969, are estimates, because some fruit were ringpicked.

3/ August 1 size is average of all blocks measured on that date. Later sizes are based on growth changes from August 1 of the blocks listed.

Table 23.--Grapefruit: Average diameter size, selected citrus blocks, Texas, 1968 and 1969 growing seasons 1/

Block	Year	Aug. 1	Sept. 1	Oct. 1	Nov. 1	Dec. 1	Jan. 1
		Inches					
Block I	1968		3.41	3.62	3.85	4.03	
	1969	2.59	2.87	3.11	3.17	3.35	3.47
Block II	1968		2.88	3.20	3.32		
	1969	2.26	2.62	2.82	3.04	3.18	3.34
I and II combined	1968		3.14	3.41	3.58		
I and II combined	1969	2.24	2.74	2.96	3.10	3.26	3.40
Special research 2/	1969	2.51		3.07	3.25	3.44	
3102	1969	2.87		3.44	3.62	3.73	
3104	1969	2.72		3.15			
3203	1969	2.91		3.57	3.72		
3210	1969	2.71		3.22			
3310	1969	2.31		2.76			
3315	1969	2.43		3.19			
3304	1969	2.40		2.96	3.13	3.27	
3306	1969	2.63		3.12	3.37	3.56	
3206	1969	2.18		2.98			

1/ Data for blocks I and II and the average of the special research blocks have been converted to equivalent first-of-month sizes.
2/ August 1 size is based on all measurements on that date. Later sizes are based on growth changes from August 1.

Increased fruit sets of grapefruit were reflected in much lower average fruit sizes. Number of fruit harvested from the sample grapefruit trees was considerably higher in 1969 than in 1968. Block I showed an increase of 141 percent and block II, 68 percent. Fruit set was doubtless even larger than the harvest comparisons, since the final visit was made 1 or 2 months later than in 1968.

Slope of the growth line for the randomly selected grapefruit blocks was slightly greater than that of the two blocks in 1968. Again, this may be due to lower average counts for the random blocks.

Average Valencia size in 1969 (both blocks combined) nearly approximated 1968 size throughout the season, but block I had larger fruit than 1968 and block II, smaller fruit. Since number of fruit harvested from each block was 6 percent higher in 1969, size differences do not reflect differences in fruit set.

Table 24.--Valencia oranges: Average diameter, selected citrus blocks, Texas, 1968 and 1969 growing seasons 1/

Block	Year	Sept. 1	Oct. 1	Nov. 1	Dec. 1	Jan. 1	Mar. 1
		Inches					
Block I	1968	2.17	2.34	2.47	2.57	2.62	2.74
	1969		2.45	2.57	2.63	2.70	2.77
Block II	1968	2.20	2.32	2.46	2.56	2.60	2.65
	1969		2.11	2.28	2.39	2.48	2.58
I and II combined	1968	2.18	2.33	2.46	2.56	2.61	2.70
I and II combined	1969		2.28	2.42	2.51	2.59	2.68
Special research 2/	1969		2.34		2.59	2.69	2.75
Block 2104	1969		2.35		2.67	2.87	2.95
Block 2202	1969		2.19		2.44	2.55	2.57
Block 2203	1969		2.21		2.39	2.56	
Block 2205	1969		2.36		2.65	2.83	2.87
Block 2206	1969		2.30		2.51	2.61	2.62
Block 2207	1969		2.59		2.79	2.87	2.90
Block 2208	1969		2.35		2.55	2.75	
Block 2210	1969		2.18		2.53	2.66	2.69
Block 2214	1969		2.32				2.71

1/ Data for blocks I and II and the special research averages have been converted to equivalent first-of-month sizes.
2/ March 1 size is based on growth change from January 1 for fruit measured in both surveys.

Average fruit size of the random Valencia blocks was slightly larger than that of the two blocks studied in 1968 and 1969. This size difference was also noted in early orange and grapefruit blocks.

Regression analyses of fruit sizes for the 1969 growing season are included in appendix A. Corresponding regression analyses for the 1968 growing season can be found in appendix D of the 1968 research reports. (See footnote 2.)

Fruit size data from the randomly selected blocks provide components of variance estimates for fruit size. Only one limb per tree was selected for fruit measurement of early oranges and grapefruit. Assumed model for fruit size was

$$Y_{ijkl} = u + c_i + b_{ij} + t_{ijk} + e_{ijkl}$$

where

Y_{ijkl} is the size of the l^{th} fruit measured on the k^{th} tree within block j of the i^{th} age category.

Analysis of variance of this model tests hypothesis of no differences between (1) trees within blocks, (2) blocks within an age class, and (3) age classes within the type of fruit. Tables 25 and 26 list the analyses of variance for early oranges and grapefruit. Average fruit sizes were 6.5792 inches in circumference for early oranges and 7.7949 inches for grapefruit.

Table 25.--Analysis of variance: Circumference, early oranges, selected citrus blocks, Texas, August 1, 1969

Source	Degrees of freedom	Sums of squares	Mean squares	F ratios	Variance components
Age class	2	12.3367	6.1683	0.6061	1/ 0
Block/age	18	183.1604	10.1755	3.2628*	.2155
Tree/block	21	65.4904	3.1185	9.3313**	.2301
Fruit/tree	558	186.4867	.3342		.3342
Total	599	447.4742			

1/ Calculated variance component is negative: a zero variance component has been substituted.

* Tabular F values at the 5-percent level are F(18,21) ≐ 2.50, F(21,558) ≐ 1.71
** Tabular F values at the 1-percent level are F(18,21) ≐ 3.35, F(21,558) ≐ 2.00

Table 26.--Analysis of variance: Circumference, grapefruit, selected citrus blocks, Texas, August 1, 1969

Source	Degrees of freedom	Sums of squares	Mean squares	F ratios	Variance components
Age class	2	92.9990	46.4995	2.0778	0.1689
Block/age	17	380.4412	22.3788	2.4876	.5556
Tree/block	20	179.9209	8.9960	20.8555**	.7625
Fruit/tree	426	183.7551	.4313		.4313
Total	465	837.1162			

* Tabular F values at the 5-percent level are $F(2,17) \doteq 4.64$, $F(17,20) \doteq 2.53$ $F(20,426) \doteq 1.71$.

** Tabular F values at the 1-percent level are $F(2,17) \doteq 7.42$, $F(17,20) \doteq 3.43$, $F(20,426) \doteq 2.00$.

The hypotheses of no differences in average early-orange size between trees within blocks and among blocks within age classes would be rejected (at the 5-percent confidence level). The hypothesis of no difference in size due to age class would not be rejected.

The hypothesis of no difference in average grapefruit size between trees within blocks would be rejected. The hypotheses of no differences in size due to blocks within ages or age classes would not be rejected.

Since the tree contribution to variance was significant for both early oranges and grapefruit, further breakdown of this component was desired. For Valencia oranges, at least two terminal limbs (on different primary limbs) were selected on each tree. Model now assumed for fruit size was:

$$Y_{ijklm} = u + c_i + b_{ij} + t_{ijk} + l_{ijkl} + c_{ijklm}$$

where

Y_{ijklm} is now the size of the m^{th} fruit on the l^{th} limb of the k^{th} tree within the j^{th} block of the i^{th} age class.

The additional hypothesis now being tested is that fruit size does not differ due to limbs within a tree. Table 27 lists the analysis of variance for Valencia fruit size. Average fruit size was 7.527 inches in circumference.

Table 27.--Analysis of variance: Circumference, Valencia oranges, selected citrus blocks, Texas, September 1, 1969.

Source	Degrees of freedom	Sums of squares	Mean squares	F ratios	Variance components
Age/class	1	25.006	25.006	1.282	0.017
Block/age	10	195.134	19.513	6.823**	.300
Tree/block	16	45.766	2.860	1.849	.055
Limb/tree	20	30.939	1.547	6.937**	.095
Fruit/limb	618	137.733	0.223		.225
Total	665	434.578			

* Tabular F values at the 5-percent level are F(1,10) = 4.96,F(10.16) ≐ 3.06, F(16,20) ≐ 1.71, F(20,618) = 1.63.
** Tabular F values at the 1-percent level are F(1,10) = 10.0,F(10.16) ≐ 4.42, F(16,20) ≐ 3.50, F(20,618) ≐ 2.00.

Table 27 indicated that the hypotheses of no differences between age classes and no differences between trees within blocks would not be rejected. The hypotheses of no differences in fruit size among blocks and between limbs within a tree would be rejected.

Analysis of the variance components indicates the number of fruit required and the type of sample needed to estimate fruit size with desired precision. It should be noted that sample size refers to estimation of fruit size on a particular date, not the forecasting of size to a later date. In the case of early oranges and grapefruit, the expression for variance of the average fruit size is

$$\sigma_{\bar{Y}}^2 = \frac{\sigma_b^2}{n} + \frac{\sigma_t^2}{na} + \frac{\sigma_f^2}{nab}$$

where

σ_b^2 is the block variance component,

σ_t^2 is the tree variance component,

σ_f^2 is the individual fruit variance component, and

n is the number of blocks,

a is trees per block,

b is the fruit per tree.

The age class component has been excluded in this example.

The number of samples needed for a particular sampling variance and for a given cost function can now be calculated as was done in optimization of sampling rates for estimating total fruit. It should be sufficient for the scope of this report to show what level of sampling variance would have resulted from a project the size of the 1969-70 research. In 1969, 21 blocks of fruit, with two trees per block and approximately 15 fruit per tree, were desired sampling rates. Using these values for n, a, b, the estimated variances of the average fruit circumferences are

$$\sigma^2_Y = \frac{2.115}{21} + \frac{.2301}{42} + \frac{.3342}{630} = 0.01627 \text{ for early oranges}$$

$$\sigma^2_Y = \frac{.5556}{21} + \frac{.7625}{42} + \frac{.4313}{630} = 0.04529 \text{ for grapefruit}$$

$$\sigma^2_Y = \frac{.300}{21} + \frac{.055}{42} + \frac{.895}{84} + \frac{.223}{1260} = 0.01690 \text{ for Valencia oranges}$$

Coefficients of variation (square root of variance divided by the mean) for the various types of fruit would be 1.94 percent for early oranges, 2.73 percent for grapefruit, and 1.73 percent for Valencia oranges. Thus, even the small sample sizes used in this research project would give coefficients of variation of the magnitude usually desired.

Optimum allocation analyses for numbers of trees per block and numbers of fruit per tree to measure were not performed. Cost estimates for the fruit size study would depend upon other operations to be carried out at the same time.

If fruit sizing is to be a separate survey, the cost of adding each block will be very high, and more trees per block and fruit per tree will have to be sampled. If a large number of blocks will be visited periodically anyway, the cost per block will be much lower and fruit sizes should be measured in more blocks.

FRUIT DROPPAGE STUDY

Indications of fruit droppage during the 1969 growing season came from both the eight continued research blocks and the new random sample of blocks. Droppage calculations were based on the limbs used for the fruit size study. Results from the previous season showed that droppage from the size measurement limbs was no higner than from a separate sample of droppage limbs.

Trees in the eight research blocks were visited monthly except February. The randomly selected blocks were not visited the month after tagging, but some were revisited later on a monthly basis until harvest.

Droppage indications are summarized in tables 28, 29, and 30. All indications are based on percentage of fruit tagged per limb. Indications for randomly selected blocks of each type are summarized as special research results.

Table 28.--Early orange droppage, selected citrus blocks, Texas, 1968 and 1969 growing seasons 1/

Block	Year	Percentage dropped between tagging and survey date			
		September 1	October 1	November 1	December 1
		Percent	Percent	Percent	Percent
Marrs I	1968	2.0	9.4	16.2	20.1
	1969	1.2	5.1	13.3	
Marrs II	1968	.8	6.1	8.8	
	1969	.0	2.4	6.2	7.3
Pineapple	1968	.6	1.9	10.2	
	1969	6.9	11.6	19.1	24.0
Jaffa	1968	1.4	4.3	9.0	
	1969	1.4	17.2	25.0	36.9
Average, 4 blocks	1968	1.2	5.4	11.0	
Average, 4 blocks	1969	2.4	9.1	15.9	22.1
Special research	1969	2/	3.1	10.3	

1/ Percentages are averages of individual limb droppage indications.
2/ Fruit were tagged about August 1 but not revisited until October 1.

Table 29.--Grapefruit droppage, selected citrus blocks, Texas, 1968 and 1969 growing seasons 1/

Block	Year	Percentage dropped between tagging and survey date				
		Sept. 1	Oct. 1	Nov. 1	Dec. 1	Jan. 1
		Percent				
Block I	1968		4.8	6.7	11.9	
	1969	0	3.8	8.1	10.8	16.7
Block II	1968		1.7	4.1		
	1969	3.8	4.9	5.6	5.6	6.3
I and II combined	1968		3.2	5.4		
I and II combined	1969	1.9	4.4	6.8	8.2	11.5
Special research	1969	2/	1.2	5.9	12.8	

1/ Grapefruit were tagged September 1, 1968, and August 1, 1969. Percentages are averages of individual limb droppage indications.

2/ Fruit were tagged about August 1 but not revisited until October 1.

Table 30.--Valencia orange droppage, selected citrus blocks, Texas, 1968 and 1969 growing seasons 1/

Block	Year	Percentage dropped between tagging and survey date				
		October	November	December	January	March
		Percent				
Block I	1968	4.5	17.6	22.2	32.1	39.0
	1969		7.9	11.0	17.6	28.1
Block II	1968	1.2	6.4	10.2	14.3	16.0
	1969		2.3	3.1	4.5	17.1
I and II combined	1968	2.8	12.0	16.2	23.2	27.5
I and II combined	1969		5.1	7.1	11.0	22.6
Special research	1969		2/	14.9	20.7	22.0

1/ Valencia oranges were tagged September 1, 1968, and October 1, 1969. Percentages are averages of individual limb droppage indications.
2/ Fruit were tagged about October 1 but not revisited until December 1.

Individual block comparisons for 1968 and 1969 show large changes in droppage patterns. These changes may be due to sampling variations or to weather. For example, loss of several fruit on the same limbs by wind damage between August 1 and October 1, 1969, increased the droppage percentages for the blocks of Pineapple and Jaffa oranges.

The total droppage of early oranges was analyzed to determine if 1969 droppage was statistically different than 1968. Results are shown in table 31.

Table 31.--Analysis of variance of early orange droppage, selected citrus blocks, Texas, 1969 growing season

Source	Degrees of freedom	Sums of squares	Mean squares	F ratios 1/
Year	1	224.855	224.855	0.712
Block/year	6	1,894.092	315.682	1.379
Tree/block	24	5,495.667	228.986	1.035
Limb/tree	64	14,159.238	221.238	
Total	95	21,773.852		

1/ Tabular F values at the 5-percent level are F(1,6) = 5.99, F(6,24) = 2.51, and F(24,64) ≐ 1.70.

None of the F-tests in table 30 are significant at the 5-percent level. That is, it is reasonable to assume that the differences between trees within blocks, between blocks within years, and between years are explained by sampling variations.

Results from the randomly selected blocks compared reasonably well with the eight block indications, at least by the end of the season. The final random indication departed greatly only for grapefruit.

The variations between years and between samples in 1969 indicate that an average droppage figure cannot be assumed. A better procedure in an operational survey would be to calculate historic parameters for an appropriate model for each month. Then current season information would be used each month to project expected droppage to expected harvest date.

REMOTE SENSING RESEARCH

Other types of research activities during the 1969-70 research project were aerial photography and an X-ray experiment.

One block of 18-year-old Valencia orange trees was selected to study the possibility of counting fruit from aerial photography. Five clusters of four trees each were selected for limb counts and photography from an aerial lift. In addition to this block, aerial photographs were taken of citrus trees at the Texas A&M Research Center and the ARS Research Farm. Actual fruit counts were made by ARS personnel for Marrs orange, Valencia orange, and grapefruit trees approximately 6 years old. Actual harvest counts were made for four grapefruit trees at the Texas A&M Center.

Main purpose of the aerial photography was to determine if fruit can be counted by this means with reasonable reliability. Other possible uses of aerial photography to be explored include:

(1) Stratification of blocks of citrus as a possible substitute or adjunct to age stratification.

(2) Canopy measurements for tree selection or fruit estimation or both.

(3) Relationship of fruit count and some optical measure of aerial photography.

Aerial photos were taken on November 19 and December 11, 1969, and on February 16, 1970, at altitudes of 500 and 1,000 feet, using aerial Ekatchrome film with a 9 1/2-inch format exposed in a Zeiss camera. First photography was blurred and probably not adequate for the counting objectives of the study. Fruit which had been placed on a white background were distorted--elongated, cucumber shaped rather than round. Motion problem appears to have been solved on the remaining flights.

Photographs from the aerial lift were taken on November 19 and 20, 1969. Shots were made at altitudes of 35, 45, and approximately 60 feet, using 70 mm Ektachrome film. Polaroid photographs were taken at the same time by an observer in the lift who counted the visible fruit and marked locations on the prints.

All fruit on one tree of each cluster of four trees were counted, and all fruit were counted on half the primary limbs of other trees. Fruit counts (actual or expanded) ranged from 333 to 1,115. The observer in the aerial lift was able to count only 4-14 fruit per tree from overhead positions and no more than 27 fruit from any oblique position. It does not seem likely that counts from photography would exceed those made from the lift unit. Counts were made from some of the aerial lift photography under 15-power magnification. Fruit counts ranged from zero to 10 from this photography--approximately half the observer counts (when any fruit could be counted at all).

The canopy area of each tree was measured from the polaroid prints. A 4-foot-square aerial marker included in each print was used to calibrate the measurements. These canopy area measurements were then correlated with estimated fruit count, trunk CSA, and primary limb total CSA, as shown in table 32.

Table 32.--Correlations of measurements of size and fruit counts, selected citrus blocks, Texas, 1969 growing season

Item	Canopy area	Fruit count	Trunk area	Primary area
Canopy area	1.00000			
Fruit count	.17217	1.00000		
Trunk area	.53391**	.06203	1.00000	
Primary area	.07442	.05069	.24785*	1.00000

* Indicates correlation differs from zero with probability = 0.95.
** Indicates correlation differs from zero with probability = 0.99.

All values in table 32 are r^2 or linear coefficient of determination values. Canopy area was better correlated with fruit count than either of the other two measures of tree size, but the relationship was not significant. This data set is not representative of normal tree size-fruit count relationships. Correlation between fruit count and tree size was negative for each of three measures of size. Some of the smallest trees in the study actually had the most fruit.

A few preliminary counts were made under 15-power magnification from the February 1970 aerial photos. Because fruit were yellow instead of green at that time, they should have been easier to detect from the air. Individual tree counts varied from 11 to 33 for the block of older Valencias and from two to 15 for the younger trees on the ARS farm. The interpreter was generally able to reproduce his counts on the same tree, with six the biggest difference between the two counts.

A very successful X-ray experiment was conducted with two potted citrus trees under laboratory conditions. Fifty-three fruit were attached to the two trees, after which a mosaic of nine 14- x 17-inch films was exposed. Counts of 54 or 55 fruit were subsequently made from the X-ray mosaic by the participants in the experiment. Personnel in the R&D Branch counted 48 fruit from a photo of the mosaic.

Development of an X-ray system which could be used under field conditions has been suggested. One proposed system would require two vehicles, one with a portable X-ray machine and the other with a fluorescent screen and wide-angle camera. Construction of the system is not advisable, however, unless exposure time can be reduced to less than the 4-1/2 minutes required in the laboratory experiment. An exposure time approaching that of a camera would be needed for field operation, since blurring due to wind movement would probably negate any X-ray results with a long exposure time.

FRUIT QUALITY RESEARCH

Samples of fruit were taken from three blocks of citrus during two sizing visits. The objective was to investigate the type of information which could be obtained on fruit quality and to examine the variability within the same block of fruit. Data obtained included BRIX reading (specific gravity), total soluble solids, acid percent, solids-to-acids ratio, and juice content.

The limited data obtained in 1969 demonstrated that some special procedures will be necessary if fruit quality development is to be studied early in the season. The procedures now followed by the inspectors are not flexible enough to determine overall quality of the crop, but only fruit within the range of their chart standards. Fruit smaller than minimum market size are normally not tested. Tree-to-tree variation within blocks was evidenced when one sample of Marrs oranges picked about October 1 failed to pass (under established picking standards), while a sample from a few rows distant did pass.

RECOMMENDATIONS

Results from this research project provide the necessary variance and cost information to design an operational objective yield survey. Three possible designs are recommended below.

First, if costs and labor supply are not limiting factors, a survey based on counting fruit on terminal limbs will give satisfactory estimates of fruit set with a modest number of blocks. A sample size of 85 blocks for each type of citrus (early oranges, Valencia oranges, and grapefruit) would give expected coefficients of variation within 10 percent at the 67-percent confidence level. This would be a definite improvement in confidence of estimates over current procedures. Sample sizes could be increased in the future if more precision was desired. A total of four fieldwork crews with supervisors would be needed during peak work periods to complete the limb counting in 7-10 days.

Within each block, limb counts should be made on two trees selected at random. Primary limb CSA's should be measured, and two primaries selected for mapping. A terminal limb selection gauge should be used for identification of all terminals on each selected primary. Two terminals on each primary should be selected for limb counts.

Fruit-size and fruit-droppage information should be collected for approximately half the sample blocks. Fruit should be tagged on each selected primary limb of the two sample trees. The fruit tagging process will give an evaluation of counting accuracy, in addition to providing the base for size and droppage studies. Size measurements should be made each month during the growing season. Two fieldwork crews would be needed for 7-10 days each month.

If labor requirements cannot be met for the limb-counting-only approach, a second approach would substitute photography for some of the limb counting. In this case, photography should be taken (one side only) of two trees in the 85 blocks of fruit. Limb counts would be made on the two photographed trees in 25-30 of the 85 blocks. The trees with both limb and photo counts would provide the yearly calibration for estimation of fruit sets.

This extensive use of photography should require only two experienced field crews instead of the four required for limb counting only. Although four to six photo-fruit counters would be needed, they should be easier to obtain (and at lower cost) than field crews. In addition, this work could be supervised and year-to-year differences reviewed. If two shifts of people counted photos each day, results would be available within 7 to 10 days as desired.

Fruit size and fruit droppage information should be obtained for each of these blocks with limb counts. One crew should be able to collect the size measurements within a 10-day period each month.

A third proposal would also provide production forecasts with smaller errors and would have even smaller workload requirement. Detailed harvest information would be obtained each year for a sample of blocks (about 100 of each type). Photography would be taken in approximately half of these blocks and limb counts in only a sample of the photography blocks. Fruit size and droppage should be collected for all limb count blocks. The harvest information would be used in a double sampling regression estimate similar to the approach outlined in this report for photo counts.

Samples of fruit should be analyzed to determine fruit maturity and fruit quality throughout the growing season. This type of information, along with data on size distribution (and expected size distribution at maturity), may prove as valuable as an improved forecast of the citrus crop.

One item which must be completed before an operational survey can be successful is an evaluation of the completeness and accuracy of the Texas Citrus Mutual listing as a sampling frame. Tree numbers have been estimated from the 1969 trunk measurement work and compared with numbers in the 1967 Mutual listing. These comparisons suggest that a large number of blocks of each type should be selected for verification of tree numbers and completeness of the frame. The blocks to be used for the field operations in the proposals above could then be selected from these blocks in the verification study. This would reduce the cost of the tree selection during the limb counting and photography portions of the project.

Verification of tree numbers can best be performed during the fall and winter months when the early orange and Valencia orange trees are more easily distinguished. Such a schedule would also provide work between the sizing and droppage visits for the field crews.

APPENDIX A

Fruit Size and Weight Regression Results

Some of the regression correlation results of the 1969 fruit size study are shown in the following tables. Regression calculations were performed on fruit measured each month of the survey and weighed at harvesttime. Similar results can be found in appendix D of the 1968 report, "Use of Photography and Other Objective Yield Procedures for Citrus Fruit," previously cited.

Average-size data in these tables might differ slightly from figures presented in the body of this report which were converted to equivalent size the first day of the month. This conversion allowed easier comparisons with 1968 data which had been similarly converted.

Size data in these tables are reported as circumference in inches. Data can be easily converted to diameter in inches if desired. Average size and standard deviation of average size can be converted to diameter equivalents by multiplying by 0.31831 (or by dividing by 3.14159). The slope of the regression line in equations predicting final size by earlier size would be unchanged by converting to diameter. However, intercept of the regression line must be multiplied by 0.31831 to convert to a diameter prediction equation.

Weight data are in grams. If an equation predicting final weight based on circumference were to be converted to a diameter basis, slope would be multiplied by 3.14159 and intercept would remain the same.

Table 33.--Regression of harvest size on August 1 size, selected citrus blocks, Texas, 1969 growing season

Block	Harvest size		August 1 size		Correlation	Regression	
	Mean	Standard deviation	Mean	Standard deviation	(r)	Slope	Intercept
Marrs I	8.21866	0.51942	5.75967	0.37444	0.66511	0.92262	2.90468
Marrs II	8.95039	.84661	6.55339	.52149	.87976	1.42824	- .40943
Pineapple	8.41243	.48954	6.65783	.40057	.79100	.96669	1.97635
Jaffa	7.92382	.66374	6.39167	.51355	.91970	1.18417	.35500
Grapefruit I	11.03400	1.36763	8.10950	.99261	.94415	1.30085	.48478
Grapefruit II	10.66393	.67871	6.93917	.61246	.88248	.97793	3.87789

Table 34.--Regression of harvest size on September 1 size, selected citrus blocks, Texas, 1969 growing season

Block	Harvest size		Sept. 1 size		Correlation	Regression	
	Mean	Standard deviation	Mean	Standard deviation	(r)	Slope	Intercept
Marrs I	8.21866	0.51942	6.91533	0.38450	0.80121	1.08236	0.73381
Marrs II	8.95039	.84661	7.52953	.64359	.96339	1.26731	- .59184
Pineapple	8.41243	.48954	7.53330	.42247	.87756	1.01688	.75198
Jaffa	7.92382	.66374	6.97314	.57283	.93315	1.08125	.38414
Grapefruit I	11.03400	1.36763	9.08425	1.09808	.96330	1.19975	.13513
Grapefruit II	10.66393	.67871	8.34434	.63775	.92809	.98769	2.42234

Table 35.--Regression of harvest size on October 1 size, selected citrus blocks, Texas, 1969 growing season

Block	Harvest size		October 1 size		Correlation	Regression	
	Mean	Standard deviation	Mean	Standard deviation	(r)	Slope	Intercept
Marrs I	8.21866	0.51942	7.51683	0.41887	0.94343	1.16989	-0.57521
Marrs II	8.95039	.84661	7.88882	.70468	.97031	1.16574	- .24591
Pineapple	8.41243	.48954	7.98000	.43208	.92302	1.04578	.06714
Jaffa	7.92382	.66374	7.49657	.58665	.93796	1.06121	- .03159
Valencia I	8.75686	.63572	7.74465	.49770	.92390	1.18011	- .38269
Valencia II	8.17429	.66113	6.52104	.56809	.83490	.97163	1.83925
Grapefruit I	11.03400	1.36763	9.78150	1.17563	.98107	1.14129	- .12948
Grapefruit II	10.66393	.67871	8.72076	.62691	.94331	1.02125	1.75783

Table 36.--Regression of harvest size on November 1 size, selected citrus blocks, Texas, 1969 growing season

Block	Harvest size: Mean	Harvest size: Standard deviation	November 1 size: Mean	November 1 size: Standard deviation	Correlation (r)	Regression: Slope	Regression: Intercept
Marrs I	8.21866	0.51942	7.95650	0.47008	0.97373	1.07593	-0.34194
Marrs II	8.95039	.84661	9.55047	.78318	.98390	1.06358	- .14375
Pineapple	8.41243	.48954	9.28600	.44962	.96754	1.05346	- .31651
Jaffa	7.92382	.66374	7.90186	.56489	.95472	1.12179	- .94037
Valencia I	8.75686	.63572	8.10546	.51143	.92998	1.15599	- .61294
Valencia II	8.17429	.66113	7.10518	.57282	.92586	1.06859	.58177
Grapefruit I	11.03400	1.36763	9.84808	1.17208	.98484	1.14915	- .28294
Grapefruit II	10.66393	.67871	9.51476	.67119	.95741	0.96812	1.45251

Table 37.--Regression of harvest size on December 1 size, selected citrus blocks, Texas, 1969 growing season

Block	Harvest size: Mean	Harvest size: Standard deviation	November 1 size: Mean	November 1 size: Standard deviation	Correlation (r)	Regression: Slope	Regression: Intercept
Valencia I	8.75686	0.63572	8.26779	0.54152	0.97111	1.14004	-0.66874
Valencia II	8.17429	.66113	7.37615	.58285	.93574	1.06141	.34518
Grapefruit I	11.03400	1.36763	10.41108	1.25759	.98804	1.07449	- .15265
Grapefruit II	10.66393	.67871	9.83214	.65528	.95894	.99322	.89843

Table 38.--Regression of harvest size on January 1 size, selected citrus blocks, Texas, 1969 growing season

Block	Harvest size: Mean	Harvest size: Standard deviation	November 1 size: Mean	November 1 size: Standard deviation	Correlation (r)	Regression: Slope	Regression: Intercept
Valencia I	8.75686	0.63572	8.65895	0.59050	0.99171	1.06765	-0.48790
Valencia II	8.17429	.66113	8.02852	.63702	.97527	1.01218	.04796

Table 39.--Regression of harvest weight on harvest size, selected citrus blocks, Texas, 1969 growing season

Block	Harvest size: Mean	Harvest size: Standard deviation	November 1 size: Mean	November 1 size: Standard deviation	Correlation (r)	Regression: Slope	Regression: Intercept
Marrs I	145.18333	26.49112	8.21866	0.51942	0.98304	50.13633	-266.87031
Marrs II	182.85826	49.42423	8.95039	.84661	.97901	57.15328	-328.68599
Pineapple	159.17391	26.97537	8.41243	.48954	.84004	46.28868	-230.22652
Jaffa	135.49019	24.85194	7.92382	.66374	.94732	35.47011	-145.56862
Valencia I	186.90695	38.93056	8.75686	.63572	.98684	60.43227	-342.28987
Valencia II	149.38518	33.42176	8.17429	.66113	.98293	49.68945	-256.79102
Grapefruit I	297.40822	62.32203	11.03400	1.36763	.90750	41.35418	-158.89368
Grapefruit II	265.32405	43.02995	10.66393	.67871	.96432	61.13798	-386.64709

Table 40.--Regression of harvest size on September 1 size and August growth, selected citrus blocks, Texas, 1969 growing season

Block	Average harvest size	Average Sept. 1 size	August growth		Correlation			Multiple regression		
			Mean	Standard deviation	r_1	r_2	R	Intercept	b_1	b_2
Marrs I	8.21866	6.91533	1.15567	0.14778	0.80121	0.39938	0.82541	0.39704	1.01039	0.72204
Marrs II	8.95039	7.52953	.97614	.23864	.96339	.67566	.96571	- .33180	1.19237	.31165
Pineapple	8.41243	7.53330	.87548	.14127	.87756	.38150	.88417	.72285	.97493	.39424
Jaffa	7.92382	6.97314	.58147	.14592	.93315	.41351	.93518	.28476	1.12248	- .32355
Grapefruit I	11.03400	9.08425	.97475	.20019	.96330	.60247	.96409	.13740	1.16389	.33194
Grapefruit II	10.66393	8.34434	1.40517	.18673	.92809	.27529	.92824	2.37613	.98246	.06390

Table 41.--Regression of harvest size on October 1 size and September growth, selected citrus blocks, Texas, 1969 growing season

Block	Average harvest size	Average Oct. 1 size	September growth		Correlation			Multiple regression		
			Mean	Standard deviation	r_1	r_2	R	Intercept	b_1	b_2
Marrs I	8.21866	7.51683	0.60150	0.17178	0.94343	0.50710	0.95443	-0.27907	1.09240	0.47602
Marrs II	8.95039	7.88882	.35929	.13605	.97031	.46847	.97157	- .40588	1.20250	- .36187
Pineapple	8.41243	7.98000	.44670	.12638	.92302	.22209	.92320	.07221	1.04110	.07227
Jaffa	7.92382	7.49657	.52343	.09337	.93796	.16840	.93902	.05113	1.07292	- .32580
Grapefruit I	11.03400	9.78150	.69725	.18239	.98107	.52413	.98240	- .10994	1.10794	.43982
Grapefruit II	10.66393	8.72076	.37641	.10407	.94331	- .00502	.94344	1.71560	1.02163	.10339

APPENDIX B

Calculation of Number of First-Stage Units

The following example illustrates the procedure used to calculate the minimum numbers of first-stage sampling units needed for a desired degree of confidence. The calculations are for an expected coefficient of variation of 10 percent at the 67-percent level of confidence. The procedure could be adjusted for any desired level of confidence.

Table 42 lists the average fruit count and variance component estimates used for optimum allocation calculations. These components are based on equal probability sampling and expansion. Data in table 42 were used in the calculation of optimum allocation values presented in table 9.

Table 42.--Average fruit per tree and variance components, selected citrus blocks, Texas, 1969 growing season

Age class	Average fruit counts	Variance components			
		Block	Tree	Primary	Terminal
Early oranges:					
0-3 years	98	26,000	5,000	12,000	20,000
4 years and over	430	37,500	35,000	84,000	318,000
Grapefruit:					
0-3 years	192	30,000	5,000	7,000	21,000
4 years and over	305	30,000	15,000	73,000	126,000
Valencia oranges:					
0-3 years	182	25,000	8,000	20,000	23,000
4 years and over	357	33,000	36,000	43,000	112,000

The example below follows the necessary steps to calculate n_h (number of samples per age class) values for early oranges. The same procedure was used for the Valencia orange and grapefruit n_h results presented in table 12.

First, the expected value of $\bar{y}$, the average count per tree is calculated:

$$\bar{y} = w_1\bar{y}_1 + w_2\bar{y}_2 = .4755\ (98) + .5245\ (430) = 272.13$$

(The w_1 and w_2 values are estimated proportion of trees in the two age classes.)

Next, the variance required for a 10-percent coefficient of variation is calculated:

$$\text{c.v.} = \frac{\sigma_{\bar{y}}}{\bar{y}} \; : \; .1 = \frac{\sigma_{\bar{y}}}{\bar{y}} \; : \; \sigma_{\bar{y}} = .1\bar{y} = .1\ (272.13) = 27.213$$

$$\sigma^2_{\bar{y}} = (27.213)^2 = 740.547$$

Thus, the variance of $\bar{y}$ desired is 740.547. The minimum number of samples required to give this variance can be found by setting equal the variance of the means of the two age classes.

That is: $\sigma_{\bar{y}}^2 = w_1^2\ \sigma_{\bar{y}_1}^2 + w_2^2\ \sigma_{\bar{y}_2}^2 = (w_1^2 + w_2^2)\ \sigma_{\bar{y}.}^2$, where $\sigma_{\bar{y}.}^2 = \sigma_{\bar{y}_1}^2 = \sigma_{\bar{y}_2}^2$

$$\sigma_{\bar{y}.}^2 = \frac{\sigma_{\bar{y}}^2}{(w_1^2 + w_2^2)} = \frac{740.547}{(.2261 + .2751)} = 1477.546$$

This value of $\sigma_{\bar{y}.}^2$ is then used to calculate the number of first stage units needed for each age class:

$$\sigma_{\bar{y}.}^2 = \frac{\sigma_b^2}{n} + \frac{\sigma_t^2}{na} + (1-f_3)\ \frac{\sigma_p^2}{nab} + (1-f_4)\ \frac{\sigma_w^2}{nabc}$$

$$= \frac{1}{n}\left[\sigma_b^2 + \frac{\sigma_t^2}{a} + (1-f_3)\ \frac{\sigma_p^2}{ab} + (1-f_4)\ \frac{\sigma_w^2}{abc}\right]$$

This formula contains only one unknown, the value of n. The values of a, b, and c are each 2, as shown in the optimum allocation section. The numbers of primary limbs per tree and terminal limbs per primary given in tables 10 and 11 are needed to calculate f_3 and f_4. The variance components are shown in table 42.

For the 0-3-years age class, the calculations are:

$$1477.546 = \frac{1}{n_1}\ 26{,}000 + \frac{5{,}000}{2} + 1-\frac{2.0}{4.6}\ \frac{12{,}000}{4} + 1-\frac{2.0}{2.9}\ \frac{20{,}000}{8}$$

$$1477.546 = \frac{30{,}912}{n_1} \; : \; n_1 = \frac{30{,}972}{1477.546} = 20.96$$

For the 4 years-and-over age class:

$$1477.546 = \frac{1}{n_2}\ [37{,}500 + 35{,}000 + 1-\frac{2.0}{9.7}\ 84{,}000 + 1-\frac{2.0}{4.5}\ 318{,}000]$$

$$1477.546 = \frac{94,857}{n_2} \quad : \quad n_2 \; \frac{94,857}{1447.546} = 64.20$$

Values of n_1 and n_2 are shown in table 12.

UNITED STATES DEPARTMENT OF AGRICULTURE
WASHINGTON, D.C. 20250

OFFICIAL BUSINESS

PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF
AGRICULTURE

U.S.MAIL