RANKING STATES FOR AREA FRAME DEVELOPMENT

Ralph V. Matthews, USDA/NASS
3251 Old Lee Highway, Rm. 504, Fairfax, VA 22030

KEY WORDS: area frame sampling, selection index, stratification

The National Agricultural Statistics Service (NASS) conducts surveys based on area frame sampling in all States except Alaska. The surveys are conducted to estimate crop and livestock production, farm costs and returns, and farm labor use. Area sampling frames are replaced in two or three States each year. Construction of a new frame costs $150,000 in a typical State with labor comprising 75 percent of the cost [2].

Since 1974, frames with systematic sampling designs have been replaced with frames having replicated sampling designs [3]. The frames in the corn-belt States were replaced in the mid 1970's, followed by the southeastern States in the late 1970's and the States west of the Rocky Mountains in the early 1980's. New frames in 1989 for West Virginia and the New England States will complete this process.

This analysis was undertaken to prioritize objectively the States which will have new area sampling frames implemented from 1991 through 1995. Objective decisions about frame replacement are in keeping with the formation of the Statistical Standards Staff (SSS) in October 1986. Selection of States for frame development was identified as an activity requiring better documentation [1]. The current analysis provides objective criteria for deciding when to replace a frame.

A decision was made when the analysis began in January 1988 that frames implemented in 1983 or later would not be considered for replacement. Data from five surveys was not considered adequate to track changes in the frames. States which are under development or planned for development and implementation through 1990 were also not considered. The 26 frames implemented from 1974 through 1982 were evaluated.

The method of ranking the States involved five indications of a State's need for a new frame. For each of the five criteria, the State most in need of a new frame received a rank of 1. The five sets of ranks were then combined into a selection index by weighting each factor's ranks, and a rank of 1 indicated the State with the greatest need for a new frame. Top-ranked States were compared for age and availability of stratification and sampling materials to make a work plan covering the last half of 1989.

The first question for evaluating the frames was, "How has the relative precision changed over the life of the frame?" The June Agricultural Survey (JAS) coefficients of variation (CV's) for the acreage of the top three crops and land in farms were examined in years 2, 3, and 4 of the frame's use versus 1985, 1986, and 1987. Starting with the second year avoided data collection problems in the first year and allowed for any sample reallocation that occurred after the first year. The average CV for the four variables in the three early years was compared with the average CV for the four variables in the three latest years, and a percentage change was calculated. An increase in the average CV indicated a frame more in need of replacement than a frame in which the average CV decreased.

The second question for examining the frames was, "How important was the area sampling frame for each State's multiple frame estimates?" Two Agency publications [4,5] presented the area frame's percentage nonoverlap contribution to the multiple frame estimates of the JAS. This was the portion of the multiple frame estimates found only by area frame sampling. The nonoverlap percent of the multiple frame variance was also calculated. The variables for 1986 were land in farms and number of farms. These two, along with cattle and calves, hogs and pigs, total cropland, and grain storage capacity were published for 1987. The percent contributions to the eight estimates and eight variances were averaged to show the reliance of the State on the area sampling frame for its nonoverlap capability. Although the nonoverlap contribution indicates more about the quality of the list frame than the area frame, a State with a high nonoverlap contribution should receive more consideration for a new area frame, because the State relies on it more.

The third question for evaluation of the frames was, "How important is the State to the national survey program?" To evaluate this, an alternative national optimum allocation based on 1987 JAS area data was examined [6]. Although this allocation was not in use operationally, it indicated the relative importance of each State. The allocation was designed for moderate improvement in the CV's of the nonoverlap portion of the multiple frame estimates. This multivariate allocation was based on two livestock, seven crop acreage, and two grain stock variables. A modification was made to the original analysis [6] to have a maximum of 1,200 segments and a minimum of 250 segments per State.

The fourth question for ranking the frames was, "What percent of a State's segments did not meet the original strata definitions?" This factor is considered routinely before stratification begins in new frame construction [7]. Ideally, the percentage would be very low, but acceptable estimates can result if it is not. The percentage was calculated using 1987 JAS area data, and the higher the percentage of segments not meeting the strata definitions, the more a State may need a new area sampling frame. Data from a single survey were deemed sufficient, because 80 percent of the sample does not change from year to year.

The fifth question for examining the frames was, "How important is a State's agriculture relative to other States?" To evaluate this, the cash receipts from livestock, crops, and government payments were averaged for 1985 and 1986 [8,9]. A ranking of the States with respect to this variable showed which State was the most important agriculturally.

The five variables were analyzed and ranked. After consultation with other statisticians of the Area Frame Section, selection index weights were chosen. A selection index was constructed by assigning weights to each variable's ranks and summing to a single index value for each State.

The average change in CV's was judged the most important variable, because change was measured over the years of the frame's use. The other four variables, based on the most recent one or two years of data, compared the frames as they are now in use. The cash receipts variable was given the lowest weight, since it is not a statistical measure of the frame's effectiveness. The nonconforming segment percentage was also given the lowest weight, because it is biased against regions with more heterogeneous land. The nonoverlap percentage contribution and the optimum allocation figures were given intermediate weights which reflected their importance relative to the other variables.

The selection index weights are shown in table 1. Two alternative sets of weights were constructed to ensure that the data and not the weights led to the final priorities. The alternatives had a 50 percent decrease or increase in the weight assigned to the average change in CV's. The difference was distributed equally among the other four weights. The first alternative put approximately equal weight on each variable's ranks, and the second alternative placed minimal weight on all but the change in CV's variable.

Table 1 -- The weights assigned to each variable's ranks in three selection indexes

Variable	Selection index weight	Lower CV weight	Higher CV weight
CV change	.40	.20	.60
NOL pct.	.20	.25	.15
Optimum alloc.	.20	.25	.15
Out-of-tolerance segments	.10	.15	.05
Cash receipts	.10	.15	.05

Table 2 contains the results for the five variables. The first variable in table 2 was the percentage change in the average of the CV's for the top three acreage crops and land in farms from years 2, 3, and 4 of the frame's use versus the average from 1985, 1986, and 1987. Since the Idaho and Texas frames were first used in 1982, the average from 1983 and 1984 was compared with the average of 1986 and 1987.

Six of the top 10 States were from the South. The average change in the CV's was greater than 30 percent in Oklahoma, Louisiana, Arkansas, and Georgia. In 15 other States the CV's increased. In seven States, the average change in the CV's was negative, and the relative precision improved during the frame's use. This variable alone would indicate that the 12 States with changes of less than 5 percent would not be candidates for new frames.

The second variable in table 2 was the nonoverlap contribution to the JAS multiple frame estimates and variances. Louisiana was in a class of its own with a nonoverlap contribution of nearly 70 percent. A new area frame will not reduce the nonoverlap contribution, because the quality of the list frame is the determining factor. Nevertheless, the States which exceeded a 50 percent nonoverlap contribution warrant more consideration for area frame replacement than the States with less than a 20 percent nonoverlap contribution.

The alternative national optimum sample allocation based on the 1987 JAS area nonoverlap domain results was the third variable in table 2. A maximum and minimum number of segments were set arbitrarily which differed from the original analysis [6]. Kansas, Texas, and Oklahoma were in a class of their own and were substantially greater than the other 23 States.

Table 2 -- Five input variables for a selection index to prioritize States for area sampling frame development

State	Avg.CV change (%)	Avg.NOL contrib. (%)	Alternative allocation[2] (n)	Nonconforming segments (%)	Cash receipts ($ billion)
ALA	14.6	45.3	290	40.0	2.11
ARK	38.8	36.0	540	36.1	3.47
CALIF	-7.8	40.6	560	25.5	14.35
GA	32.2	39.9	395	35.7	3.36
IDAHO	2.1[1]	38.3	395	24.2	2.17
ILL	17.9	21.5	380	35.0	8.01
IND	-1.8	29.9	370	33.9	4.67
KANS	3.7	20.9	1,200	41.0	6.26
KY	12.5	32.4	250	30.2	2.74
LA	48.6	69.9	460	37.4	1.55
MINN	10.4	28.4	440	32.9	6.91
MISS	1.3	48.5	620	35.2	2.16
N Y	-3.6	36.0	250	52.3	2.61
N C	8.3	33.4	250	34.9	3.91
N DAK	-13.9	19.8	250	31.8	3.11
OHIO	16.8	39.7	550	35.8	3.98
OKLA	52.4	47.3	855	44.0	2.96
OREG	-12.9	34.2	375	28.0	1.88
PA	13.4	41.8	250	46.3	3.19
S C	4.9	54.2	290	40.9	1.01
S DAK	1.0	19.4	540	51.0	3.03
TENN	20.2	36.4	280	38.7	2.07
TEX	6.4[1]	46.2	1,200	37.1	9.78
VA	17.2	43.5	295	33.6	1.65
WASH	-2.7	50.5	530	27.5	3.02
WIS	-2.0	32.6	250	38.8	5.29

[1] 1983,1984 vs. 1986,1987
[2] The 1,200 maximum and 250 minimum differ from the original analysis [6]

Table 3 -- Rankings of five input variables for a selection index to prioritize States for area sampling frame development

State	Avg.CV change	Avg.NOL contrib.	Alternative allocation	Nonconforming segments	Cash receipts
	ranks				
ALA	9	7	18	7	21
ARK	3	15	7	12	10
CALIF	24	10	5	25	1
GA	4	11	12	14	11
IDAHO	17	13	12	26	19
ILL	6	23	14	16	3
IND	20	21	16	18	7
KANS	16	24	1	5	5
KY	11	20	21	22	17
LA	2	1	10	10	25
MINN	12	22	11	20	4
MISS	18	4	4	15	20
N Y	23	15	21	1	18
N C	13	18	21	17	9
N DAK	26	25	21	21	13
OHIO	8	12	6	13	8
OKLA	1	5	3	4	16
OREG	25	17	15	23	23
PA	10	9	21	3	12
S C	15	2	18	6	26
S DAK	19	26	7	2	14
TENN	5	14	20	9	22
TEX	14	6	1	11	2
VA	7	8	17	19	24
WASH	22	3	9	24	15
WIS	21	19	21	8	6

The fourth variable in table 2 was each State's percentage of segments not conforming to the sampling frame strata definitions. The 1987 JAS area results showed that in New York and South Dakota more than 50 percent of the segments did not meet the strata definitions. In all 26 States, the figure exceeded 24 percent.

The fifth variable in table 2 was the average cash receipts from livestock, crops, and government payments in 1985 and 1986. The $4 billion figure divided the group of 26 States after rank 7. The top seven States were California, Texas, and five midwestern States.

Table 3 contains the rankings for the variables in table 2. A number 1 ranking indicates the frame most in need of replacement for each variable. For example, Oklahoma's average CV change of 52.4 percent ranked number 1 for that variable in table 2. The ranks in table 3 were multiplied by the weighting factors to construct the selection index and its two alternatives. Table 3 facilitates comparisons between the ranking of each variable and the ranking of the selection index values.

Table 4 contains the ranks for the selection index and its two alternatives. Based on the actual selection index values, a top-ranked group of 15 States was identified for further evaluation. The robustness of the selection index weights was evident, because each State had similar ranking for the three sets of weights. This occurred when the variables were weighted approximately equally and when the change in CV's variable was dominant.

The selection index rankings were most like the rankings for the change in CV's variable. Oklahoma, Louisiana, and Arkansas were ranked in the same order by both methods. Although Texas ranked 14 for changes in CV's, rankings of 1 for the optimum allocation and 2 for agricultural receipts raised it to a rank of 4 for the selection index. Tennessee and Illinois ranked 5 and 6, respectively, for the changes in CV's variable, but low rankings for the nonoverlap contribution and optimum allocation variables caused them to drop 5 (Tennessee) and 3 (Illinois) positions in the selection index ranking.

In table 5, the top-ranked group of 15 States is presented with rankings based on availability of current satellite images, aerial photographs, and topographic maps.

Table 4 -- Selection index values to prioritize States for area sampling frame development

State	Selection index	Lower CV weight ranks	Higher CV weight ranks
ALA	7	10	10
ARK	3	5	3
CALIF	17	11	21
GA	5	6	4
IDAHO	20	21	18
ILL	9	14	7
IND	23	23	22
KANS	13	8	15
KY	21	24	16
LA	2	3	2
MINN	15	17	12
MISS	12	7	14
N Y	24	22	24
N C	18	19	17
N DAK	26	26	26
OHIO	6	4	5
OKLA	1	1	1
OREG	25	25	25
PA	8	9	11
S C	14	12	13
S DAK	19	18	20
TENN	10	16	6
TEX	4	2	7
VA	11	15	7
WASH	16	13	19
WIS	22	20	23

Table 5 -- Fifteen States ranked by selection index and availability of current stratification and sampling materials

State	Selection index rank	Materials rank
OKLA	1	13
LA	2	4
ARK	3	11
TEX	4	14
GA	5	10
OHIO	6	12
ALA	7	2
PA	8	8
ILL	9	9
TENN	10	7
VA	11	6
MISS	12	3
KANS	13	1
S C	14	5
MINN	15	15

Based on the two rankings, Louisiana and Alabama will receive new frames in 1991. When Alabama's frame construction is underway, the availability of materials will be updated. The other 13 top-ranked States will be evaluated for availability of materials prior to completion of Alabama's frame. This analysis should be repeated within 4 years to order the next set of States.

Table 6 shows one result of the objective ranking process. These frames are at least 10 years old, but frame replacement will not be considered at this time. Chronologic age would have dictated that these States be considered, but the statistical measures in this analysis showed that the frames were performing satisfactorily and should not be replaced.

Table 6 -- States with area sampling frames at least 10 years old and their selection index ranks

State	Age (years)	Selection index rank
IND	12	22
KY	11	21
N C	10	19
N DAK	11	26
S DAK	12	18
WIS	11	23

This analysis ranked 26 States with objective criteria to determine the order in which States receive new area sampling frames. Data were drawn from published Agency estimates, publications, and research reports. Fifteen top-ranked States were then evaluated for availability of current stratification materials. Louisiana and Alabama will be the first two States to receive new frames as a result of this analysis.

1. Brannen, Bill, and Mike Craig. State of Standards Survey. U. S. Dept. of Agriculture, National Agricultural Statistics Service, April 1987.

2. Cotter, Jim, and Jack Nealon. Area Frame Design for Agricultural Surveys. U. S. Dept. of Agriculture, National Agricultural Statistics Service, Aug. 1987.

3. Fecso, Ron, Robert D. Tortora, and Frederic A. Vogel. Sampling Frames for Agriculture in the United States. Journal of Official Statistics, 1986, 2:279-292.

4. Hill, George W., Roberta B. Pense, and Mike A. Steiner. List Frame Evaluation from an Area Frame Perspective -- Book II. U. S. Dept. of Agriculture, National Agricultural Statistics Service, March 1987.

5. Holland, Theresa E., Dwaine C. Nelson, and Raymond R. Bosecker. List Frame Evaluation from an Area Frame Perspective -- Book III. U. S. Dept. of Agriculture, National Agricultural Statistics Service, Aug. 1987.

6. Mergerson, James W. "National Area Sample Reallocation Analysis" Research Forum presentation. U. S. Dept. of Agriculture, National Agricultural Statistics Service, June 3, 1988.

7. National Agricultural Statistics Service. Area Frame Analysis Package. U. S. Dept. of Agriculture, National Agricultural Statistics Service, revised June 1984.

8. U. S. Dept. of Agriculture. Agricultural Statistics 1986.

9. U. S. Dept. of Agriculture. Agricultural Statistics 1987.