United States Department of Agriculture

National Agricultural Statistics Service

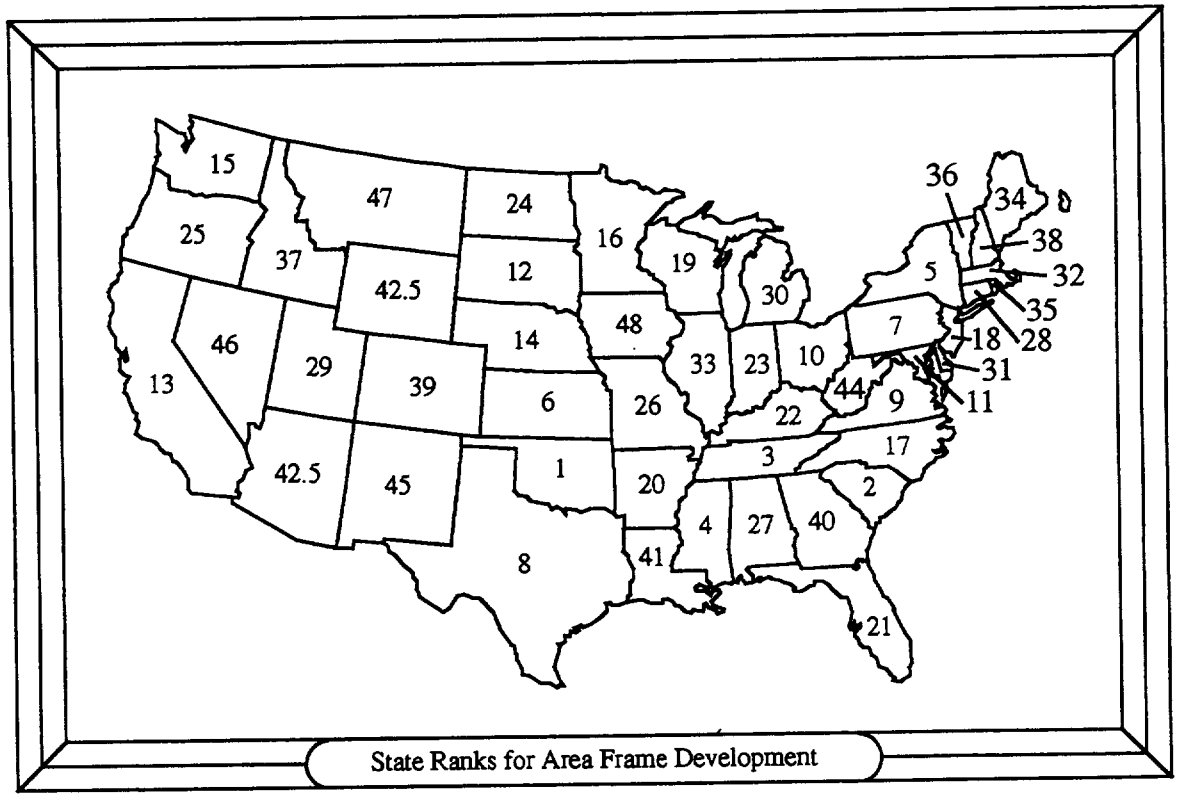
Survey Management Division

NASS Staff Report Number SSB-93-01

January 1993

Ranking the States for Area Frame Development

Jeffrey D. Bush



State Ranks for Area Frame Development

RANKING THE STATES FOR AREA FRAME DEVELOPMENT, by Jeffrey D. Bush, Survey Management Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250, January 1993. Staff Report No. SSB-93-01.

ABSTRACT

Area sampling frames play an important role in the National Agricultural Statistics Service's estimating program. Criteria are developed which capture the current condition of each state area sampling frame. Through weighting of the criteria, an overall ordering of the states for new frame construction is established. New York is chosen to receive a new frame in June 1994 along with California. New frames for South Carolina and Kansas will enter in June 1995 and a new Tennessee frame is scheduled for deployment in June 1996.

KEY WORDS: area sampling frame, stratification, sample allocation

This paper was prepared for limited distribution.

ACKNOWLEDGMENTS

The author would like to thank Bob Hale, Carol House, Gail Hanneman, and Cathy Tomczak of the Survey Sampling Branch, and Ron Bosecker of the Research Division for their comments, suggestions, and support.

RANKING THE STATES FOR AREA FRAME DEVELOPMENT

Jeffrey D. Bush

BACKGROUND

The Area Frame Section constructs and manages area sampling frames for every state except Alaska. Probability surveys actuate the frames in the generation of estimates for crop and livestock production, farm costs and returns, and farm labor. More commonly, the area frame is used to supplement an incomplete list frame in what is known as a multiple frame survey.

The utilization of land within each state is forever changing. As a result, states' area frames contain an increasing number of segments over time that do not conform to their stratum's definition. This occurrence, in turn, damages a frame's ability to produce useful and accurate estimates. Frames exhibiting this characteristic are said to be "aging". Since the Area Frame Section is only able to construct two new frames per year, a systematic approach to prioritize the states for new frame development is used.

Ralph Matthews' report, Ranking States for Area Frame Development, outlines the 1988 effort to prioritize the states for area frame development. The approach consisted of: 1) Deciding upon criteria, or standards, by which to judge each frame, 2) Ranking the states for each individual criteria, 3) Assigning weights, or relative importance, to each criteria, and 4) Using the weighted ranks to arrive at an overall ordering based on all criteria. This paper employs a similar approach, using 1991 JAS data, with the following exceptions.

The current analysis includes every state's area frame as opposed to the previous work which excluded every frame younger than six years old along with frames currently under development. This decision had much to do with the percentage change in coefficient of variation criteria as "Data from five surveys was not considered adequate to track changes in the frames. (Matthews, 1988)" Inclusion of all states in the analysis, however, acts as a quality control mechanism. The complete analysis, data, and conclusions, gathered at one point in time, will also be useful in future analyses.

The major criteria in the Matthew's report was the percentage change in the coefficients of variation for the top three crops in each state. Coefficients of variation for the top three crops were averaged in the second, third, and fourth year of each frame. To allow for any data collection problems or sample reallocation, data from the first year of each frame was not used. A percentage change was then produced by dividing this number with an average of the same coefficients of variation in the last three years of the frame. Frames which yielded a higher percentage change were said to be more in need of a new sampling frame. This criteria is excluded from the current analysis for several reasons.

As many area sampling frames currently in use were constructed in the mid to late seventies, major crop coefficients of variation for all states are next to impossible to locate. Further, it may be the case that major crops within a state have changed over an extended period of time. Finally, unless a reason is known for the deterioration of the major crop coefficients of variation, construction of a new frame may not improve the situation at all.

Another notable difference between the two analyses is the inclusion of a variable representing the agricultural receipts of each state in the previous analysis. The problem being that it includes the cash receipts of crops important to the national estimating program along with many other high priced rare commodities. This extraneous information damages the effectiveness of the measure. A national level optimal allocation study, based on important crops at the national level, is described later as a better estimation of the relative importance of states.

CRITERIA

The current analysis begins with a summary of the criteria employed. Each criteria description contains the basis for its inclusion in the analysis, the statistic used to capture the criteria in numerical terms, the derivation of the statistic, and a table of the results.

1) Evaluation of the quality by which strata partition the sample in each frame.

Basis:

Stratification is the division of a population into homogeneous groups using a variable (with known values) that is highly correlated with what is to be estimated. In terms of NASS area frames, this variable is the percentage of cultivation associated with a particular unit of land. Assuming primary sampling units within each stratum vary little from one another, it is reasonable to conclude that a precise estimate of a commodity can be obtained with a relatively small sample. However, as the variability between primary sampling units in the same stratum grows, increasingly larger samples are required to match the previous precision. This increasing variability, caused by changing land usage, is frequently the primary factor in the deterioration of major commodity expansions.

In order to get a general idea of how well segments within each stratum cluster together, a box plot of strata verses percent of cultivation was produced for the sample in each state. The strata were collapsed according to the ten's digit of their label to maintain comparability across all states. For example, Florida's strata 13, 17, and 18 were collapsed into an intensively cultivated stratum typically defined as land which is greater than 50 percent cultivated. Strata 21, 22, and 27 were collapsed into an extensively cultivated stratum which usually embodies land of which 15 to 50 percent is dedicated to agriculture. The urban, range, and non-agricultural strata were constructed in a similar fashion. Dashed lines appear where usual strata breaks occur.

Figure 1 compares a box plot from a frame thought to be "aging" to a frame where segments are stratified efficiently. You will note that the majority of segments in New York's intensively cultivated stratum fall well below the usual 50 percent cultivated stratum break. More importantly, serious overlap is present between the intensive and extensive strata. For there is no distinct difference in the percentage of cultivation between the segments in each stratum. Iowa's frame, on the other hand, may take advantage of increased gains in precision due to its exceptional stratification.

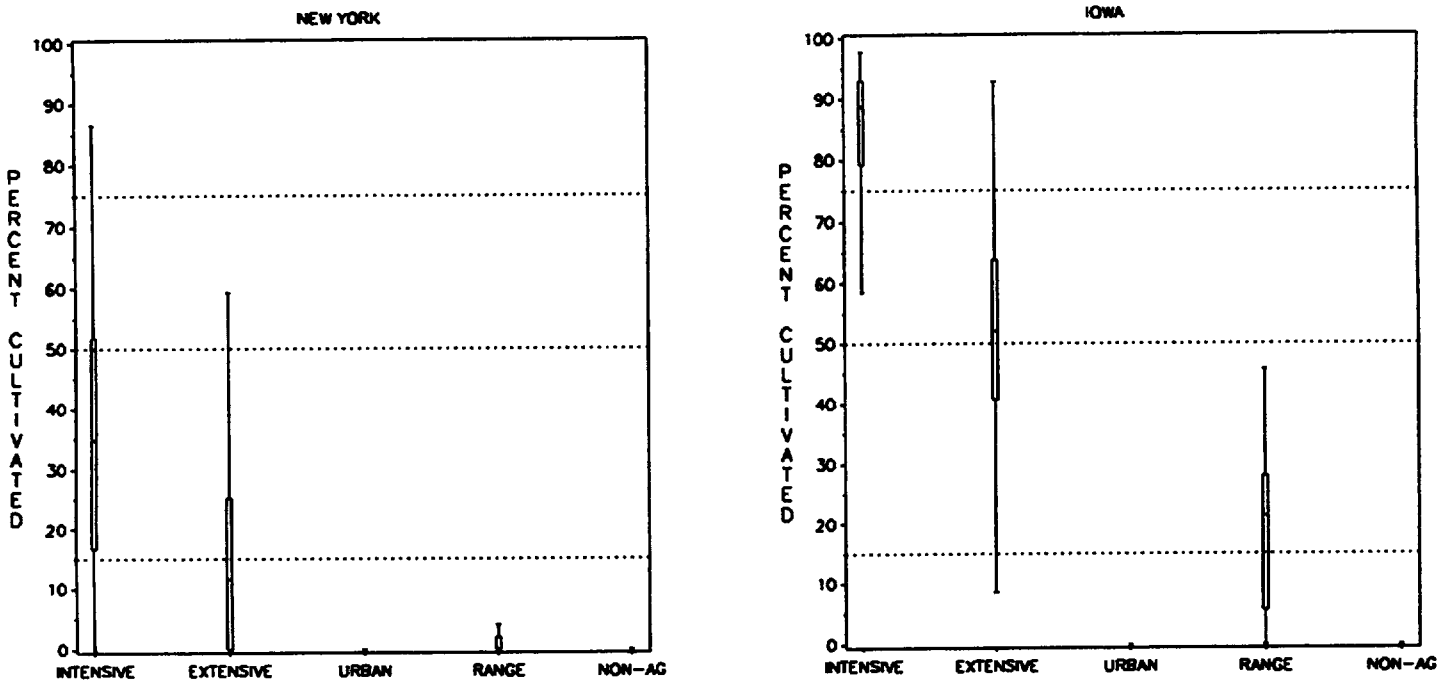


Figure 1: Distribution of 1991 Segments Using Collapsed Strata

The percentage of segments currently meeting their stratum's definition is used as a numerical measure to approximate the results of the box plots.

Statistic:

Percentage of Segments Not Meeting their Stratum's Definition

Derivation:

A segment level data set was produced for each state which included a variable indicating whether the percentage of cultivation within each segment fell within the bounds of its stratum definition. The total number of nonconforming segments divided by the adjusted sample size in each respective state yielded the percentage of nonconforming segments at the state level. These percentages were ranked from highest to lowest. Segments belonging to strata not utilizing a percent cultivated definition, such as strata 31, 32, and 50, were not included in the calculation.

5) Availability of current photography.

Basis:

The CASS system eliminates the need for most paper maps and photography in the stratification and sample select processes. However, segments are still outlined on photography from the ASCS centers for enumerator usage. For the inability of an enumerator to correctly identify a chosen segment contributes additional non-sampling error to the estimates.

Statistic:

Year of new NAPP photography availability minus the present year 1992.

Table 5: Availability of Current Photography

Rank	NAPP Acquisition Dates (1992)	Availability	States
1	1995	1992, 1997	Connecticut, Iowa, Massachusetts, Missouri, New Jersey, New York, North Dakota, Oklahoma, Rhode Island, South Carolina
2	1996	1993, 1998	Kansas, Maine, Minnesota, Mississippi, Montana, New Hampshire, South Dakota, Texas, Vermont, Washington
3	1992	1989, 1994	Alabama, Arizona, Delaware, Idaho, Indiana, Maryland, Michigan, New Mexico, Pennsylvania, Tennessee, Wisconsin
4	1993	1990, 1995	Colorado, Georgia, Illinois, Kentucky, Nebraska, North Carolina, Ohio, Utah
5	1994	1991, 1996	Arkansas, California, Florida, Louisiana, Nevada, Oregon, Virginia, West Virginia, Wyoming

Derivation:

The National Aerial Photography Program circulates an acquisition plan each year which associates future flight dates with each state. A goal of the program is to provide complete coverage of the nation every five years. Generally, complete coverage of a state may not be purchased for two years after it has been flown.

6) Age of each sampling frame.

Basis:

Each frame was ranked based on the number of years the frame has been in use. A state's framework, residing on the original paper maps, must be queried each year when the sample select and sample preparation units prepare rotation segments. Generally, the older a state's frame, the more difficult it is to efficiently manage its sample.

Statistic/Derivation:

The current year (1992) minus the first year of implementation for each frame. The first year of each frame may be found in the current Area Frame Design Information report.

WEIGHTING THE CRITERIA

Next, weights must be assigned to each criteria in order to arrive at an overall rank for each state. This is undeniably the most subjective part of the process as the importance of each of the criteria must be judged. Ultimately, three weighting schemes will be developed and a final ordering will ensue.

Weights were first assigned to each criteria according to how well they captured the present condition of each frame. For example, the percentage of segments not meeting their stratum's definition was given the highest weight as this was judged to have a major effect on a frame's current performance. Lesser weights were assigned to criteria which helped to further distinguish problem frames in terms of the availability of materials and age. The other criteria were assigned weights which fell between the two extremes.

Table 6 presents the three weighting schemes applied to the criteria. The weighting scheme just described is labeled A in Table 6. Weighting schemes B and C are variations on A. In weighting scheme B, criteria 1 and 2 were given less weight while the other criteria were boosted. Weighting scheme C increased the weights assigned to criteria 1 and 2 while all other weights were lowered.

Table 6: Weighting Schemes

Criteria	A	B	C
1. Percentage of segments not meeting their stratum's definition	0.300	0.200	0.400
2. Average of the NOL percent of the total multiple frame expansions for Number of Farms, Land in Farms, and Cropland	0.200	0.150	0.250
3. Output from the multivariate allocation program	0.150	0.200	0.100
4. Storage medium of each frame.			
5. Number of years until new NAPP photography availability.	0.100	0.125	0.075
6. Age of frame.			

RESULTS

Table 7 includes the top ranked states and statistics with respect to the three weighting schemes. States will be selected for new frame construction from this table until another analysis is completed. It can be seen that Oklahoma ranked first with respect to all three weighting schemes. This occurrence verifies the self checking mechanism of the analysis as Oklahoma also ranked

first in the Matthews report and was reconstructed in 1992. California, which is currently under construction, ranked 13th, 15th, and 14th for weighting schemes A, B, and C respectively. The fact that California was the last state to use a point sampled stratum weighed heavily in its selection. The concluding paragraphs cover the selection of the next four states to receive new area sampling frames. They are followed by a complete listing of state ranks in Table 8

Due to time constraints, the next state to receive a new area sampling frame had to be chosen before the completion of this report. The selection of New York was based primarily upon the percentage of segments not meeting the current stratum definitions in that state. In addition, a meeting held February 13, 1990 to outline "Priorities for New Area Frames" revealed many complaints about misclassified land. New York's box plot presented in Figure 1 also supports the selection.

Table 7: Summary of Candidates for New Frame Construction

State	Ranks for Three Schemes			Statistics					
	A	B	C	% Non Conforming Segments	NOL % of Total Rank	National to Current Allocation	Storage Medium	NAPP Photo Availability	Age of Frame
Oklahoma	1	1	1	49.1	8	Below	Plan	1992	18
South Carolina	2	2	2	46.0	13	Below	Plan	1992	13
Tennessee	3	4	3	45.2	10	Below	Plan	1994	15
Mississippi	4	5	4	41.6	11	Below	Plan	1993	14
New York	5	6	7	61.4	39	Below	Plan	1992	13
Kansas	6	3	11	42.9	43	Above	Plan	1993	17
Pennsylvania	7	8	5	51.1	3	Below	Plan	1994	11
Texas	8	7	9	36.0	23	Above	Altek	1993	10

South Carolina, Tennessee, and Mississippi appear after Oklahoma in Table 7. Over 40 percent of the segments in these frames are non-conforming, the multiple frame relies heavily on the area frames, none are digitized, and all are over 12 years old. South Carolina will be constructed after New York's completion since it ranked second in all three schemes and complete NAPP coverage was made available in 1992.

It is desirable to avoid the concurrent entrance of two new area frames in the same region of the country because of potential impact on regional estimates. In accordance with this policy, Kansas will follow New York and South Carolina as it is member of the North Central region of the country. Apart from its high instance of non-conforming segments, the national optimal allocation yielded evidence that an increased sample size may be necessary to reach national precision goals.

Tennessee will receive a new area sampling frame after Kansas for reasons stated previously. Complete NAPP coverage should become available in 1994 or the approximate time of its construction. ■

Table 8: State Ranks With Respect to Three Weighting Schemes.

State	A	B	C
Alabama	27.0	30	21.5
Arizona	42.5	39	44.0
Arkansas	20.0	25	13.0
California	13.0	15	14.0
Colorado	39.0	37	41.0
Connecticut	28.0	28	27.0
Delaware	31.0	29	32.0
Florida	21.0	26	16.0
Georgia	40.0	44	35.0
Idaho	37.0	33	39.0
Illinois	33.0	27	40.0
Indiana	23.0	20	25.0
Iowa	48.0	47	48.0
Kansas	6.0	3	11.0
Kentucky	22.0	21	24.0
Louisiana	41.0	46	36.0
Maine	34.0	35	31.0
Maryland	11.0	19	6.0
Massachusetts	32.0	32	33.0
Michigan	30.0	38	28.0
Minnesota	16.0	12	20.0
Mississippi	4.0	5	4.0
Missouri	26.0	24	26.0
Montana	47.0	42	47.0
Nebraska	14.0	14	15.0
Nevada	46.0	45	45.0
New Hampshire	38.0	40	38.0
New Jersey	18.0	22	17.0
New Mexico	45.0	43	46.0
New York	5.0	6	7.0
North Carolina	17.0	17	18.0
North Dakota	24.0	16	29.0
Ohio	10.0	10	10.0
Oklahoma	1.0	1	1.0
Oregon	25.0	23	23.0
Pennsylvania	7.0	8	5.0
Rhode Island	35.0	34	37.0
South Carolina	2.0	2	2.0
South Dakota	12.0	9	12.0
Tennessee	3.0	4	3.0
Texas	8.0	7	9.0
Utah	29.0	31	30.0
Vermont	36.0	36	34.0
Virginia	9.0	11	8.0
Washington	15.0	13	19.0
West Virginia	44.0	48	43.0
Wisconsin	19.0	18	21.5
Wyoming	42.5	41	42.0

REFERENCES

Matthews, Ralph V. (1988). Ranking States for Area Frame Development. United States Department of Agriculture, National Agricultural Statistics Service, Research and Applications Division, NASS Staff Report SSB Number 88-08.

Mergerson, James W. (1989). Area Frame Sampling: Sample Allocation. United States Department of Agricultural, National Agricultural Statistics Service, Research and Applications Division, Internal Documentation.

National Agricultural Statistics Service (1989). Standards for Target Coefficients of Variation for Major Probability Surveys. Policy and Standards Memorandum Number 45-89.

Scott, Carl, Rex Patterson, and Jeff Geuder. 1991 NASS List Frame Evaluation. United States Department of Agriculture, National Agricultural Statistics Service, Research and Applications Division, NASS Staff Report SSB Number 92-01.