United States
Department of
Agriculture

Statistical
Research
Service

Statistical
Research
Division

SRS Staff Report
Number YRB-85-01

April 1985

1983 Soybean Objective Yield Nonsampling Error Research Study

Robert Battaglia

1983 SOYBEAN OBJECTIVE YIELD NONSAMPLING ERROR RESEARCH STUDY. By Robert J. Battaglia, Statistical Research Division; Statistical Reporting Service; U.S. Department of Agriculture; Washington, D.C. 20250; April 1985. Staff Report No. YRB-85-01.

ABSTRACT

Changes in training, questionnaire design and quality control procedures for collecting soybean objective yield data can reduce the level of nonsampling errors. This exploratory study identifies sources of nonsampling errors in soybean objective yield data collection and determines if narrow-row soybeans could be sampled using broadcast units. Data were collected during 1983 in Minnesota and Ohio. Results showed differences in number of plants and per plant component counts which may suggest inclusion biases and enumerator fatigue. Broadcast units should not be used in narrow-row soybeans and different sampling procedures should be investigated for narrow-row soybeans.

KEYWORDS

Soybean Objective Yield, inclusion bias, multivariate paired-t, narrow-row soybeans, enumerator fatigue.

ACKNOWLEDGMENT The author thanks Radley Edwards, SRS, for assistance with questionnaire design and the enumerator manual. The Minnesota and Ohio SSO's are to be given credit for data collection and editing. Finally, the author thanks Bessie Johnson for typing the paper.

Washington, D.C. April 1985

CONTENTS

SUMMARY

Changes in training, questionnaire design, and quality control procedures for collecting soybean objective yield data can reduce the level of nonsampling errors. The Statistical Reporting Service (SRS), USDA, studied soybean objective yield data collection procedures in 1983. The purposes of the study were to identify sources of nonsampling errors during data collection that could affect yield forecasts and to investigate the possibility of using "broadcast" sampling units in narrow-row soybeans. A research sample unit was laid out in addition to operational units in Minnesota and Ohio. The major findings were:

1. Constructing a broadcast unit using row 1 as a side can cause an upward bias in the number of rows included in the unit.

2. Comparing data from operational and research units, in both States, showed no differences in per plant counts made in the 6-inch sections. This allows inference of plant count analysis on research units to operational units. Comparing the number of plants per 6-inch section showed that Ohio research 6-inch sections contained significantly more plants than the operational units. This was probably the result of an inclusion bias. There was no evidence of bias in Minnesota data.

3. Comparing per plant counts showed that counts from narrow-row units were less than counts from wide-row units. Wide-row units contained an average of four plants per row in the 6-inch sections while narrow-row units averaged less than two plants per row. If a soybean objective yield unit contains no plants in the 6-inch section, a yield forecast was computed by substituting State average values. For narrow-row soybeans, this would overestimate true yield and underestimate yields for wide-row units.

4. Comparing per plant counts from the "first" plant to an average count over all plants in the 6-inch section (operational) showed significant differences during the August and September surveys. The plant component counts from the first plant were larger. This indicates that fatigue and difficult enumeration conditions affect data collection. Counts of blooms and pods on laterals showed the largest differences.

Based on these findings, we recommend the following:

1. Train survey statisticians and enumerators on the effects that biases and nonsampling errors have on forecasts and estimates.

2. Do not use broadcast units in narrow-row soybeans where rows can be distinguished.

3. Further investigate the use of a larger detailed count area in narrow-row soybeans.

4. Change the design of Form B questionnaire to reduce nonsampling errors:

   a. Record counts from 6-inch sections by plant. (Some States are doing this already)

   b. Combine counts of blooms and pods on laterals and main stem into one question. These counts are combined for use in forecast models.

   c. Discontinue counts of lateral branches after August in Northern States and after September in Southern States. These counts are not used in forecast models after this time.

   d. Use State recommended row width and seeding rates to edit number of plants per unit.

   e. Change edit of Form B questionnaire to account for differences in plant component counts between narrow- and wide-row soybeans.

1983 SOYBEAN OBJECTIVE YIELD NONSAMPLING ERROR RESEARCH STUDY.

Robert J. Battaglia

INTRODUCTION

A soybean objective yield research study was conducted in Minnesota and Ohio in 1983 by the Yield Research Branch of the Statistical Reporting Service (SRS), USDA. The purposes of this study were to identify sources of nonsampling errors during data collection that could affect the forecasts of pods per plant, and to investigate the possibility of using a "broadcast" unit in narrow-row fields. Results of the analysis indicate that changes in questionnaire design and quality control procedures could reduce the present level of nonsampling errors. Also, broadcast units should not be set out in narrow-row fields due to a row inclusion bias.

The objective yield forecast of gross yield per acre is determined by multiplying three values--forecasted number of plants, forecasted number of pods with beans per plant, and a historical average weight of beans per pod. Forecasts of number of pods with beans per plant are made using multiple regression equations which use detailed counts from a 6-inch by two row plot as the independent variables (9)[1/]. The plant components counted are number of plants, mainstem nodes, blooms and dried flowers and pods, pods with beans, and lateral branches (7). The number of plants squared and the number of pods with beans squared are also used as independent variables. Heavy vegetation, unfavorable field and weather conditions, and fatigue can make these counts difficult for enumerators.

Studies have shown that enumerator behavior, training, questionnaire design, and interview length can all result in response errors (2, 5, 10). In many cases, however, the enumerator is unjustly blamed as the primary source of error.

Soybean destructive counting studies attempted to reduce possible counting errors by removing from the field those plants used for detailed counts (4, 6). The counts were then made in more comfortable surroundings. Nelson found significant differences in

1/ Numbers in parentheses refer to literature cited at the end of this report.

counts of nodes and pods in two States over the analysis period. The destructive counts were higher than the operational counts indicating the enumeration conditions may have affected counts. Handling damage from previous visits also affected the operational counts. The destructive method was not adopted because of forecast performance problems.

The overall objective of this exploratory study was to determine methods to reduce the amount of nonsampling error in the current soybean objective yield program. This was done by comparing counts recorded on individual plants with counts determined by the current method of aggregating over all plants in the 6-inch section.

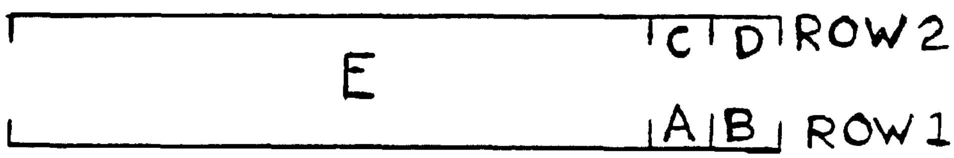
DATA COLLECTION

Data was collected in Minnesota and Ohio during the 1983 Soybean Objective Yield Survey. A research sample unit was located in each sampled field 10 rows and 10 paces from the second operational soybean unit. Two types of research units were used, wide-row and narrow-row. If the field row width was less than 1.5 feet, a narrow-row research unit was laid out. Otherwise a wide-row research unit was used. Enumerators used 1.5 feet as a criterion to decide whether a unit was a wide- or narrow-row unit based on extension service definitions of narrow-rows. Also, in the operational objective yield program enumerators are instructed to enter a row width of 1.5 feet for broadcast units. Identifiable rows were needed to construct the research units. If no rows could be identified, the enumerator was instructed to construct a broadcast research unit. Since there were only four broadcast observations, this type of unit will not be discussed, but an example can be found in the interviewer's manual (8).

Wide-Row Units

The plant and plant component counts made on the wide-row research units were the same as those made on regular objective yield units. The wide-row research unit was identical to the operational sample unit plus it included a second 6-inch count unit (see figure 1).

Figure 1: Wide-Row Research Unit



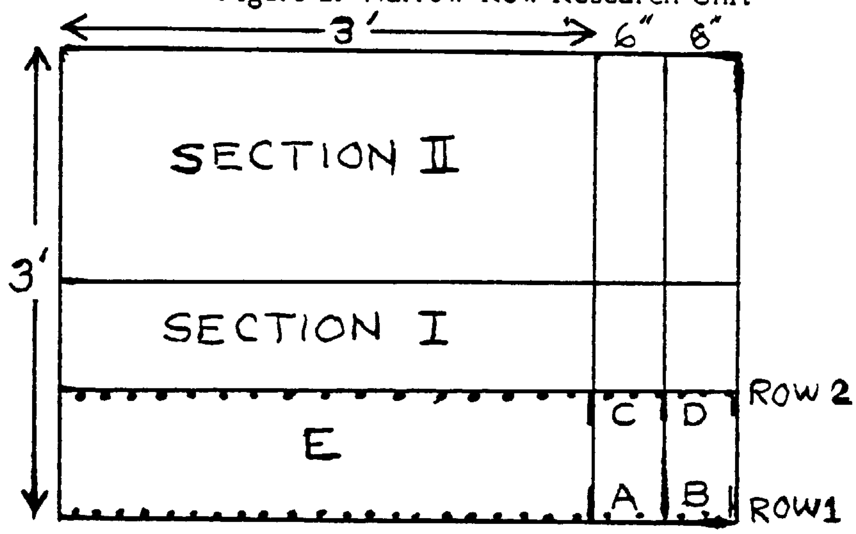
A First 6-inch count unit, row 1
B Second 6-inch count unit, row 1
C First 6-inch count unit, row 2
D Second 6-inch count unit, row 2.
E 3-foot sections, row 1 & 2

Another basic difference between the research and the operational sample units was the method of recording plant component counts (6-inch counts) on the questionnaire. On the regular objective yield questionnaires, the counts made in 6-inch row sections are totaled over all plants, then recorded. On the research questionnaire, counts on the first four plants closest to the 3-foot section were recorded separately for each row, along with the location of each plant within the first or second 6-inch count section. Counts for any remaining plants were totaled, then recorded by count section (see questionnaire, Appendix I).

Narrow-Row Units

The narrow-row research unit was used when the average row width was less than 1.5 feet (figure 2).

Figure 2: Narrow-Row Research Unit



This unit was constructed like a broadcast unit with an additional 6-inch count unit. It was also designed to include measurements from a wide-row research unit and an operational unit for comparison (see figure 2, rows 1 & 2). Note that any expansions of plant numbers using total plants in section I and II would contain an upward bias because the unit is constructed using row 1 as a boundary. This construction tends to maximize the number of rows in the unit. An adjustment for bias can be made by multiplying the average number of plants per row by a ratio of frame length (3 foot) over one row space (see Appendix II). We based the analysis of plant counts in this study on data from row 1 and row 2 (identical to wide-row units) which would not be affected by the bias. Counts of plant components which are per plant counts would not be affected either.

Survey statisticians in Minnesota and Ohio edited the questionnaires using the SRS Generalized Edit System. Edit limits were identical to those of the regular objective yield edit.

ANALYSIS AND RESULTS

Comparison of Operational vs. Research Counts

The first step of the analysis procedure was to compare data from the operational soybean objective yield units with identical data from the research unit to see if counts between research and operational plots were different. If the data were not different, then results from analysis in the research unit would apply to the operational units. Research and operational data were paired from each sampled field to reduce the effects of between field variation on the comparisons. A univariate t-test was used to compare differences in plant counts between operational and research units while a multivariate paired t-test was used on the five plant component variables (3). The multivariate technique was used since several measurements, which may be correlated, are made on the plants in the 6-inch sections. The two tailed hypotheses used for these univariate and multivariate paired t-tests are as follows:

H_o: operational counts = research counts.
H_a: operational counts ≠ research counts.

Table 1 shows the paired mean differences by State between operational and research units for the number of plants per first and second 6-inch sections. For the operational counts, plant numbers were averaged over four rows in two units; for the research units, plant numbers were averaged over rows 1 and 2 (see figures 1 and 2).

The t-statistic presented in table 1 is univariate since only a comparison of plants per 6-inch section was made. Table 1 shows significant differences in plant numbers between Ohio operational and research wide-row units during September. The mean differences indicate that the Ohio research 6-inch sections contained more plants than the regular soybean units in the same field. A test comparing the operational 6-inch plant counts to plant counts in the second 6-inch research section gave similar results. Counts for Minnesota units were not significantly different.

Table 1--Mean differences in plants per 6-inch section operational minus research units, September 1983, Soybean Objective Yield

State	6-inch section	Row type2/	n	Mean diff	T	Pr < |T|
Minn	First	N	15	-.23	-.35	.733
	First	W	83	-.12	-.03	.978
	Second	N	13	-1.27	-1.62	.132
	Second	W	80	-.41	-1.09	.277
Ohio	First	N	44	.10	.35	.725
	First	W	58	-.97	-2.36	.022**
	Second	N	34	.60	-1.37	.181
	Second	W	58	-.93	-1.98	.053*

1/ Operational counts were averaged over four rows in two units while research counts were averaged over two rows.
2/ N = Narrow-row units, Row width < 1.5 foot.
W = Wide-row units, Row width $\geq$ 1.5 foot.
3/ * = t statistic significant at $\alpha = .10$
** = t statistic significant at $\alpha = .05$

The results suggest possible inclusion or location biases for wide-row research units in Ohio which would affect the number of plants in the sample unit. The procedure to construct the wide-row research unit was identical to the operational procedure except that a second 6-inch section was constructed. An inclusion bias would occur if an enumerator included an extra plant in the 6-inch section when laying out the unit. Extra plants in the sample unit would affect the estimate of number of plants per acre which would give an upward bias to yield per acre. A location bias could also result in extra plants in the research unit. This would have happened if research units were located where the plants were "better" than others in the field. Enumerators were instructed to lay out research units 10 rows and 10 paces from the second operational unit using operational procedures. The effects of both these biases can be reduced by training and alerting enumerators to the consequences.

A multivariate paired T-test was used to compare differences in per plant counts from the operational and first research 6-inch section. The plant components counted in the 6-inch section were nodes, lateral branches, blooms and pods on laterals, blooms and pods on mainstem, and pods with beans. These counts are used in regression models to forecast final numbers of pods with beans per plant. In the multivariate case, the test of hypothesis (mentioned previously) is for a vector of mean differences. The multivariate test showed no difference in per plant counts between operational and research units for any State, month, or row width type. There was a difference in plant numbers rather than plant component counts. The results suggest that an inclusion bias rather than a location bias was responsible for the difference in plant numbers in the Ohio wide-row research units.

Analysis Within Research Unit

This section shows differences in plant numbers and per plant counts between wide- and narrow-row units. Analysis focused on plant counts in the first 6-inch section of the research unit. These counts are comparable to operational counts. Counts on the first four plants in each row closest to the 3-foot section were recorded separately. The location of each plant was also recorded (first or second 6-inch section). Counts for remaining plants were totaled, then recorded by location (see questionnaire Appendix I). The mean counts for the first 6-inch section of the research unit are presented in Appendix III. The narrow-row per plant counts were generally smaller than the counts from wide-row units. Also, the coefficients of variation (CV) of the means were higher for the counts from narrow-row units. The CV expresses the standard error as a percent of the mean. In the current soybean objective yield summary system, if a soybean unit contains no plants in the 6-inch section a State average is used to forecast pods per plant. For narrow-row units, this practice causes an upward bias in yield forecasts since these units have lower per plant counts. Yield forecasts from wide-row units are downward biased.

Planting narrow-row soybeans has become more common in many States, especially Ohio. There is concern in these States that a 6-inch count section is too short for narrow-row soybeans. Recommended seeding rates in Ohio are 2.4 seeds per foot of row in 7-inch rows, 4.6 seeds per foot of row in 14-inch rows, and 6.1 seeds per foot in 20-inch rows (1). Narrow-row 6-inch sections contain an average of 1.5 plants per row or three plants per unit to make detailed counts on (see Appendix III). Wide-row units contain an average of four plants per row in the 6-inch section or eight plants used to obtain per plant counts.

Table 2 shows the probabilities of plants 1 through 4 being included in the first 6-inch section for wide- and narrow-row units in Ohio for September data. The plant number indicates its position relative to the 3-foot section with plant 1 being closest to the 3-foot section. These are the four plants whose counts were recorded individually on the research questionnaire. For wide-row units, a 6-inch count unit is sufficient with the probability of a fourth plant 62-percent. In the narrow-row units, plant 1 had a 68-percent probability of being in the first six inch section (32 percent of narrow-row first 6-inch sections contained no plants). The probabilities of the second through fourth plants suggest that a 6-inch count unit is too small for narrow-row beans. The above seeding rates, less plants per 6-inch section, and higher CV's for plant counts from narrow-row units support evaluation of a longer detailed count unit for narrow-row soybeans.

Table 2--Probability of plants included in the first 6-inch section Ohio data, September 1983, soybean objective yield

Plant[1/]	Wide-row probability	Narrow-row probability
Plant 1	.98	.68
Plant 2	.90	.39
Plant 3	.77	.21
Plant 4	.62	.13

1/ Plant 1 is the plant closest to the 3-foot section with plant 2 being the second plant from the 3-foot section.

Comparison of Plants Within Row

The final part of the analysis attempted to measure the effect of enumerator fatigue on the per plant counts made in the 6-inch sections. This was done by comparing counts from the first plant in row 1 to all other plants in row 1. Counts from plant 1 were compared with plant 2, plant 3, plant 4, average counts from all plants in the first 6-inch section (operational), and average counts from all plants in the second 6-inch section. A comparison was not made when either group contained no plants. We assumed that the counts on plant 1 would be the most accurate since this plant was counted first. The effects of fatigue would then lower counts from succeeding plants. We used a multivariate paired t-test to compare counts of nodes, laterals, blooms and pods on laterals, blooms and pods on mainstem, and pods with beans. The hypothesis used for this test is as follows:

H_o: Plant 1 counts = other row 1 plant counts.
H_a: Plant 1 counts > other row 1 plant counts.

Multivariate t^2 statistics were computed for wide-row and narrow-row units in Minnesota and Ohio for August, September, and October data. The mean differences of plant 1 counts minus the other plant counts in row 1 are presented in tables 3-5. If the assumption that counts from plant 1 are no different than those from the other plants is true, then the mean differences should statistically be near zero. Due to the large number of comparisons (15 each month) there is a 54-percent chance that one or more of the comparisons would be erroneously declared significant at the .05 significance level. Even with this problem, the data in tables 3-5 are useful in an exploratory study to show which components are difficult to enumerate.

Table 3 shows the data for August. Only even numbered samples were laid out in this month so the number of observations was reduced. The small number of observations made mean differences for narrow-row units in both States undetectable. For Minnesota wide-row units, there was a significant difference between plant 1 counts and per plant counts from the first and second 6-inch sections. The mean differences show plant 1 having an average of four more blooms per lateral than the per plant counts from the 6-inch section. August data from Ohio wide-row units showed positive mean differences for all comparisons indicating that counts were being missed on other plants in the row. There was a significant difference in per plant counts between the first and fourth plants and between the first plant and the first 6-inch section counts. The largest mean difference values occurred for blooms and pods per lateral and blooms and pods on the mainstem.

The mean differences for the September survey are listed in table 4. The narrow-row data for Minnesota was again disregarded due to the low number of observations. In Minnesota, wide-row unit counts were again significantly different between the first plant and the first and second 6-inch section counts per plant. The Ohio narrow-row units showed significant differences between plant 1 and the second 6-inch section. Wide-row units in Ohio showed significant differences in counts between the first and fourth plant and both 6-inch sections. Again the mean differences for all comparisons were positive in Ohio with the largest values associated with counts of blooms and pods.

Table 5 presents October data. Minnesota wide-row units show a significant difference between plant 1 counts and first 6-inch section counts. Ohio narrow-row units showed a significant difference between plant 1 counts and counts in the second 6-inch section. The absolute values of the mean differences in October were generally smaller than the other two months, probably because of easier enumeration conditions. By October the leaves have begun to fall off the plants, most pods contain beans, lateral branches have been tagged on previous visits, and counts of nodes are no longer made.

Over three months, 10 of the 12 significant differences occurred with comparison of counts from the first plant to those accumulated over a 6-inch section. Even with the problem of multiple tests affecting significance due to type I error, the pattern of significant differences associated with counts over 6-inch sections suggests a problem with that method of recording data.

In summary, results suggest that enumeration conditions, counting methods, and fatigue can affect the per plant counts. These counts are used in the forecast models to predict the number of pods with beans per plant. During August and September the effects of condition, etc., were more prevalent in counts recorded over the 6-inch sections. In wide-row units the mean differences with the largest absolute values in both States were associated with counts of blooms and pods on laterals and main stems. These are the most difficult counts for the enumerator to make. Tables 3-5 also show both positive and negative mean differences while the mean differences for Ohio were almost always positive. In Ohio, the counts for the first plant were always greater than counts for all other plants in the 6-inch section of row 1.

Table 3--Mean differences, August data
plant 1, row 1 counts minus other row 1 counts
1983 Soybean Objective Yield

State	Row[1/] type	Comparison[2/] (plant 1-)	n	Nodes	Lats	Blooms lats	Blooms main	Pods w beans	Sig[3/]
Minn[4/]	W	Plant 2	43	-.02	.12	1.47	-2.05	--[5/]	
	W	Plant 3	42	-.38	-.07	-.98	-1.76	--	
	W	Plant 4	41	.66	.29	2.93	2.27	--	
	W	1st 6"	34	-.25	-.05	4.02	-1.88	--	**
	W	2nd 6"	33	-.09	-.19	4.76	.58	--	*
Ohio	N	Plant 2	20	-.60	.05	-.45	-1.50	--	
	N	Plant 3	12	.33	.25	.42	3.17	--	
	N	Plant 4	9	.44	.22	1.78	7.11	--	
	N	1st 6"	8	.02	.44	6.25	-.21	--	
	N	2nd 6"	9	.06	-.10	3.01	-1.47	--	
	W	Plant 2	26	.62	.42	3.92	2.00	--	
	W	Plant 3	25	.68	.40	2.08	2.80	--	
	W	Plant 4	23	.35	.57	4.83	5.13	--	*
	W	1st 6"	14	1.27	.69	6.92	7.23	--	**
	W	2nd 6"	10	1.52	.89	10.57	1.26	--	

1/ N = Narrow-row units, row width < 1.5 foot; W = Wide-row, row width $\geq$ 1.5 foot.
2/ Plant 1 is closest plant to 3-foot section in Row 1, first 6-inch is average over all plants in 6-inch section adjacent to 3-foot section (operational), second 6-inch is all plants in second 6-inch section.
3/ * = t^2 statistic significant at $\alpha = .10$, ** = t^2 statistics significant at $\alpha = .05$.
4/ Minnesota narrow-row differences excluded due to low number of observations.
5/ Counts of pods with beans are not made in August.

Table 4.--Mean differences, September data
plant 1, row 1 counts minus other row 1 counts
1983 Soybean Objective Yield

State	Row 1/ type	Comparison 2/ (plant 1-)	n	Nodes	Lats	Blooms lats	Blooms main	Pods w beans	Sig 3/
Minn 4/	W	Plant 2	77	-.22	.21	3.03	.94	2.16	
	W	Plant 3	76	-.30	-.22	-2.01	-.08	-1.57	
	W	Plant 4	73	.03	.10	1.36	1.41	.03	
	W	1st 6"	75	-.35	-.11	5.40	.57	-1.14	**
	W	2nd 6"	70	-.48	-.15	5.85	.66	-2.73	**
Ohio	N	Plant 2	30	.43	0	.60	4.07	1.13	
	N	Plant 3	19	1.97	.32	2.32	10.84	6.84	
	N	Plant 4	14	2.36	.93	6.79	12.93	7.21	
	N	1st 6"	23	.40	.22	7.38	3.28	2.26	
	N	2nd 6"	19	-.26	-.21	11.08	1.21	-1.10	**
	W	Plant 2	58	.53	.03	.85	1.02	1.50	
	W	Plant 3	55	.31	.26	2.42	1.31	2.25	
	W	Plant 4	53	1.15	.43	3.40	5.26	5.85	*
	W	1st 6"	50	.50	.21	5.44	1.69	2.43	**
	W	2nd 6"	49	.38	.30	6.60	1.62	2.19	**

1/ N = Narrow-row units, row width $<$ 1.5 foot; W = Wide-row, row width $\geq$ 1.5 foot.
2/ Plant 1 is closest plant to 3-foot section in Row 1, first 6-inch is average over all plants in 6-inch section adjacent to 3-foot section (operational), second 6-inch is all plants in second 6-inch section.
3/ * = t^2 statistic significant at $\alpha = .10$, ** = t^2 statistics significant at $\alpha = .05$.
4/ Minnesota narrow-row differences excluded due to low number of observations.

Table 5.--Mean differences, October data
plant 1, row 1 counts minus other row 1 counts
1983 Soybean Objective Yield

State	Row[1/] type	Comparison[2/] (plant 1-)	n	Nodes	Lats	Blooms lats	Blooms main	Pods w beans	Sig[3/]
Minn[4/]	W	Plant 2	73	--[5/]	.26	1.84	-.12	1.64	
	W	Plant 3	72	--	-.28	-1.35	-.93	-2.31	
	W	Plant 4	70	--	.01	.29	-.24	-1.50	
	W	1st 6"	71	--	-.10	2.45	-1.05	-1.88	**
	W	2nd 6"	66	--	-.18	3.11	-.87	-1.88	
Ohio	N	Plant 2	29	--	-.07	-.48	1.79	1.97	
	N	Plant 3	18	--	.17	1.17	5.22	-6.11	
	N	Plant 4	12	--	.83	3.50	2.50	5.92	
	N	1st 6"	21	--	.19	4.90	1.23	2.44	
	N	2nd 6"	20	--	-.41	3.41	-1.20	-1.56	*
	W	Plant 2	58	--	-.03	.31	.41	1.57	
	W	Plant 3	56	--	.23	1.39	1.11	2.98	
	W	Plant 4	53	--	.21	1.40	2.23	3.70	
	W	1st 6"	51	--	.12	2.43	.48	1.61	
	W	2nd 6"	55	--	.01	2.17	1.28	1.64	

1/ N = Narrow-row units, row width < 1.5 foot; W = Wide-row, row width ≥ 1.5 foot.
2/ Plant 1 is closest plant to 3-foot section in Row 1, first 6-inch is average over all plants in 6-inch section adjacent to 3-foot section (operational), second 6-inch is all plants in second 6-inch section.
3/ * = t^2 statistic significant at α = .10, ** = t^2 statistics significant at α = .05.
4/ Minnesota narrow-row differences excluded due to low number of observations.
5/ Counts of nodes are not made in August.

CONCLUSIONS

The analysis has illustrated some potential problems with soybean objective yield procedures. These problems can be corrected by improving training and simplifying some data collection procedures.

First, broadcast units should not be used in "drilled" soybeans where the enumerator can identify rows. Using row 1 as a boundary for the broadcast unit will result in a row inclusion bias (Appendix II). This bias will cause an overestimate in the number of plants per 18 square feet, thus, affecting the yield. There is no procedure at present to randomly locate the boundary of the broadcast unit between the rows. An adjustment can be made for bias when row measurements are possible by multiplying the average number of plants per row by a ratio of frame length to row width. Broadcast units should only be used when rows cannot be distinguished.

The comparison of per plant component counts between research units and operational units, in the same field, showed no significant differences. There was a significant difference in the number of plants per 6-inch section between operational and research wide-row units in Ohio. This difference was probably the result of an inclusion bias caused when Ohio enumerators laid out the research 6-inch sections. Enumerators need more training on the cause and effects of biases.

An analysis of means and frequencies from the research data showed that the per plant counts from narrow-row units are smaller than those from the wide-row units. If a soybean objective yield unit contains no plants in the 6-inch section, a yield forecast is computed by substituting State average per plant values. For narrow-row soybeans this could cause an overestimate of true yield. Wide-row yields would be underestimated. The summary system should be adjusted to substitute average values based on row types. Wide-row units contained approximately four plants per row in the 6-inch sections, while narrow-row units contained two plants per row. The CV's for the per plant counts from the narrow-row units are higher than those from the wide-row units. The higher CV's show a need for more plants in detailed count sections for narrow-row beans. The recommended planting rates for Ohio narrow-row and the location of plants over the 6-inch sections suggest that a longer count unit would be appropriate for drilled soybeans.

Comparisons of plant counts within the 6-inch section of row 1 in the research unit showed significant differences between counts from the first plant and counts averaged over all plants in the first 6-inch section (operational). The significant differences were found in Minnesota and Ohio for wide-row units during the August and September surveys. Wide-row units contained more plants per 6-inch section. Also, enumeration conditions are difficult during August and September. Counts of blooms and pods on laterals showed the largest mean differences. The signs of the differences for blooms and pods per lateral were always positive indicating counts on the first plant were always larger. This result suggests fatigue affects the counts.

Finally, the mean differences for all but one comparison in Ohio wide-row units were positive. All measurements on the first plant counted were always larger than those made on succeeding plants. This effect was not present in the Minnesota data.

REFERENCES

1. Beuerlein and Walker, Ohio Soybean Performance Trials 1982, Cooperative Extension Service, Ohio State University, Agronomy Department Series 212, AGDEX 141, 1983.

2. Cannell and Oksenberg, New Questionnaire Design Techniques for Reducing Response Errors, Survey Research Center, University of Michigan, Proceedings of the 44th Session of the International Statistical Institute, Madrid 1983.

3. Kramer, Clyde Y., A First Course in Methods of Multivariate Analysis, Virginia Polytechnic Institute and State University, Blacksburg, VA, 1972.

4. Nelson, D.C., Soybean Objective Yield Destructive Counting Study, U.S. Department of Agriculture, Economics and Statistics Service, Staff Report No. AGESS801218, Dec. 1980.

5. Noelle-Neumann, E., Dullness and Monotony as Problems of Questionnaire Methodology, University of Mainz FRG, Proceedings of the 44th Session of the International Statistical Institute, Madrid, 1983.

6. Pense, Roberta B., 1980 Soybean Destructive Counting Study, U.S. Department of Agriculture, Statistical Reporting Service, Staff Report No. AGESS810915, Sept. 1981.

7. U.S. Department of Agriculture, Statistical Reporting Service, 1983 Soybean Objective Yield Survey - Enumerators Manual, Washington, D.C. 1983.

8. U.S. Department of Agriculture, Statistical Reporting Service, 1983 Minnesota and Ohio Soybean Research Study - Enumerator Manual, Washington, D.C. 1983.

9. U.S. Department of Agriculture, Statistical Reporting Service, Objective Yield Survey -Supervising and Editing Manual, Washington, D.C. 1983.

10. Zarkovich, S.S., Quality of Statistical Data, Food and Agriculture Organization of the United Nations, Rome, 1966.

APPENDIX I Research Questionnaire

The following questionnaire was used for narrow-row units. The questionnaire used for wide-row units is identical except that page 3 is not needed.

UNITED STATES DEPARTMENT OF AGRICULTURE
STATISTICAL REPORTING SERVICE

Form Approved
O.M.B. Number 0535-0088

C.E. 12-00348-1A
Minnesota and Ohio

FORM B-2A: SOYBEAN YIELD COUNTS
September 1, 1983
Broadcast and rows 18 inches or greater

YEAR,CROP,FORM,MONTH (1—4)	
323X2	

RESEARCH UNIT LOCATION

	UNIT 2
Number of rows	+ 10
Number of paces	+ 10

Planting Pattern Code from Item 6, Form AA	369
Date (__________)	370
Starting Time	371

ROW SPACE MEASUREMENTS

2. a. Measure distance from stalks in Row 1 to stalks in Row 2 .... Feet & Tenths [301 .]

   b. Measure distance from stalks in Row 1 to stalks in Row 5 .... Feet & Tenths [304 .]

OBSERVATIONS WITHIN 3-FOOT UNITS

	Row 1	Row 2
3. Number of plants in row	306	307

4. Stage of Maturity. *Circle maturity code:*

Pods Forming or Earlier	Pods Set Leaves Still Green	Pods Filled Leaves Turning Yellow	Pods Turning Color Leaves Shedding	Pods Brown Almost Mature	Pods Mature
300 1	300 2	300 3	300 4	300 5	300 6
If unit is in Stage 1, 2, 3 or 4 skip to Item 6.				*If unit is in Stage 5 or 6 only, complete Item 5.*	

5. When maturity is in Stage 5 or 6, strip all pods *(all sizes with or without beans)* from all plants in Row 1 of the 3-foot unit. Also pick up all beans in Row 1 middle. Do not count pods or beans that are stripped. Deposit the pods and beans in a paper bag. Attach an ID tag and mail to the Regional Lab. *Be sure to make the counts for the 6-inch row sections on the back of this form.*

FORM B-1B: SOYBEANS (Cont'd)

<u>Counts for 6-inch row sections in front of unit</u>

SECTION I

If no plants are present, enter dashes (—) for items 6 through 11. Complete all items for Row 1 before starting on Row 2.

Starting Time [310]

	First Row						Second Row					
	Plant 1	Plant 2	Plant 3	Plant 4	Remaining Plants 1	Remaining Plants 2	Plant 1	Plant 2	Plant 3	Plant 4	Remaining Plants 1	Remaining Plants 2
Enter Code .... 1 = *First six-inch Section* 2 = Second six-inch Section	200	201	202	203	204	205	206	207	208	209	210	211
6. Number of PLANTS....................	220	221	222	223	224	225	226	227	228	229	230	231
7. Number of NODES on main stem of plants..............................	240	241	242	243	244	245	246	247	248	249	250	251
8. Number of LATERAL BRANCHES with blooms, dried flowers, or pods........... *(Attach pink tags to those lateral branches with blooms, dried flowers, or pods.)*	260	261	262	263	264	265	266	267	268	269	270	271
9. Number of BLOOMS, DRIED FLOWERS and PODS on lateral branches...........	280	281	282	283	284	285	286	287	288	289	290	291
10. Number of BLOOMS, DRIED FLOWERS and PODS on main stem...............	400	401	402	403	404	405	406	407	408	409	410	411
11. Number of PODS with beans *(include all pods in which beans have begun to form) For any row, if item 11 is greater than (item 9 + 10) recount items 9, 10 and 11.*	420	421	422	423	424	425	426	427	428	429	430	431
Ending Time	311				313	315	317				319	321

FORM B-1B: SOYBEANS (Cont'd)

	SECTION I Remaining Plants		SECTION II Plant 1	Plant 2	Plant 3	Plant 4	Remaining Plants	
Enter Code .... 1 = First six-inch Section; 2 = Second six-inch Section	212 1	213 2	214	215	216	217	218 1	219 2
6. Number of PLANTS	232	233	234	235	236	237	238	239
7. Number of NODES on main stem of plants	252	253	254	255	256	257	258	259
8. Number of LATERAL BRANCHES with blooms, dried flowers, or pods .... *(Attach pink tags to those lateral branches with blooms, dried flowers, or pods.)*	272	273	274	275	276	277	278	279
9. Number of BLOOMS, DRIED FLOWERS and PODS on lateral branches	292	293	294	295	296	297	298	299
10. Number of BLOOMS, DRIED FLOWERS and PODS on main stem	412	413	414	415	416	417	418	419
11. Number of PODS with beans *(include all pods in which beans have begun to form)* *For any row, if item 11 is greater than (item 9 + 10) recount items 9, 10 and 11.*	432	433	434	435	436	437	438	439
Ending Time	323	325	327				329	331

12. Did you pick ANY pods or beans from Row 1 and send them to the Regional Lab? .... YES ☐ NO ☐

Unit Status 380

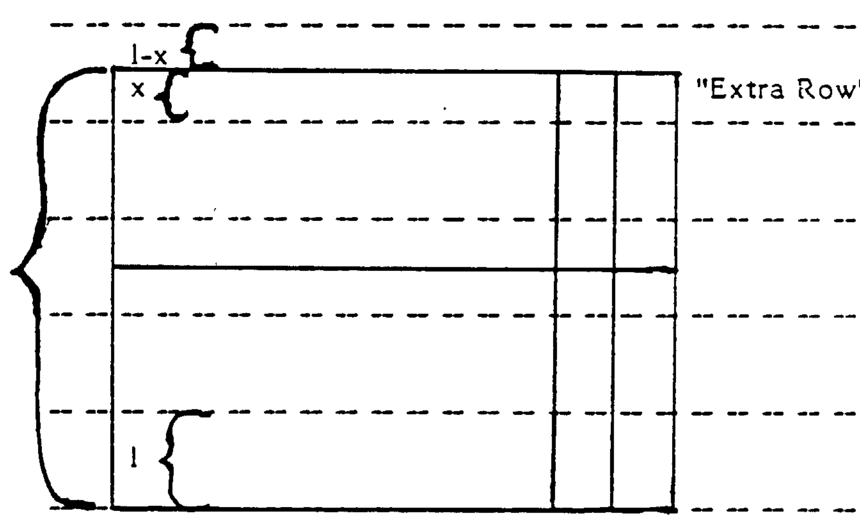
Enumerator Number 390

Supervisor Number 391

Enumerator ______________________________

APPENDIX II Adjustment for Row Inclusion Bias in Narrow-Row Units[1]

Counts from the narrow-row research unit can be expanded to an acre basis using the area of the unit. Since this unit is constructed using Row 1 as a side there is a built-in row inclusion bias. This bias maximizes the number of rows included in the sample unit. To be properly located a starting point would have to be randomly located in the row middle but this was not procedurally possible. However the bias for total number of plants could be adjusted for by multiplying average plants per row by a ratio of frame length to row width. This is explained below:



r	=	Number of row middles.
l	=	Row length.
L	=	Frame length (3 ft.).
x	=	$L - rl$.
$r+1$	=	Number of rows in unit.
T_r	=	Total plants in r rows.
T_e	=	Total plants in extra row.
$E(T)$	=	Expected value of total plants.
x/l	=	Probability of extra row in sample area.
$(l-x)/l$	=	Probability of extra row not in sample area.

$$E(T) = x/l\,(T_r + T_e) + ((l-x)/l)T_r$$
$$= (xT_r + xT_e)/l + (lT_r - xT_r)/l$$
$$= (lT_r + T_e)/l$$
$$= T_r + (x/l)\,T_e$$

$$\text{Avg. plants per row} = \sum^{r+1} T_{ri} / (r+1)$$

$$T_r = (r \sum T_{ri}) / (r+1)$$

$$E(T) = (r\,\Sigma T_{ri}) / (r+1) + x/l\,(\Sigma T_{ri}/r+1)$$

$$= \frac{lr + x}{l} \frac{\Sigma T_{ri}}{r+1}$$

$$= \frac{L}{l} \frac{\Sigma T_{ri}}{r+1}$$

$$\text{Adjustment factor} = \frac{\text{Frame length (3 ft.)}}{\text{1 Row space}} \times \text{Average plant/row}$$

1/ Developed by Ben Klugh, Yield Research Branch in an unpublished note.

Mean per plant counts in first 6-inch section
row 1 and row 2, research unit[1/]

COUNT	Month	State	Row type [2/]	n	Mean	Std. error	CV[3/]
Nodes per plant	Aug	Minn	N	4	11.48	1.24	.11
			W	43	11.59	.38	.03
		Ohio	N	19	10.80	.64	.06
			W	26	11.84	.53	.04
	Sept	Minn	N	10	13,89	.44	.03
			W	81	14.03	.29	.02
		Ohio	N	29	12.96	.47	.04
			W	56	14.76	.35	.02
Laterals per plant	Aug	Minn	N	4	1.50	.89	.59
			W	43	.99	.12	.12
		Ohio	N	19	.88	.34	.37
			W	26	.78	.19	.24
	Sept	Minn	N	10	1.52	.36	.24
			W	81	1.56	.11	.07
		Ohio	N	26	1.16	.17	.15
			W	56	1.47	.15	.10
	Oct	Minn	N	10	1.37	.28	.20
			W	77	1.50	.11	.07
		Ohio	N	28	1.42	.22	.15
			W	55	1.45	.14	.10
Blooms & pods per laterals	Aug	Minn	N	3	6.00	2.71	.45
			W	37	6.40	.62	.10
		Ohio	N	10	5.55	.73	.13
			W	16	7.08	.79	.11
	Sept	Minn	N	9	7.60	1.62	.21
			W	81	8.09	.36	.04
		Ohio	N	26	5.70	.60	.11
			W	53	6.42	.45	.07
	Oct	Minn	N	9	4.17	.89	.21
			W	76	5.10	.27	.05
		Ohio	N	27	2.60	.38	.15
			W	53	3.51	.30	.09
Blooms & pods per main	Aug	Minn	N	4	15.06	5.71	.38
			W	43	17.21	1.83	.11
		Ohio	N	19	11.91	2.62	.22
			W	26	20.83	3.09	.15

COUNT	Month	State	Row type 2/	n	Mean	Std. error	CV3/
	Sept	Minn	N	10	24.84	2.79	.11
			W	81	30.17	1.22	.04
		Ohio	N	29	30.53	2.53	.08
			W	56	33.30	1.42	.04
	Oct	Minn	N	10	12.92	1.90	.15
			W	77	19.88	.93	.05
		Ohio	N	28	16.64	2.53	.15
			W	55	19.62	1.36	.07
Pods w/beans per plant	Aug	Minn	N	4	0	0	--
			W	43	.94	.41	.44
		Ohio	N	19	0	0	--
			W	26	.12	.12	1.00
	Sept	Minn	N	10	18.43	3.29	.18
			W	81	27.67	1.64	.06
		Ohio	N	29	18.25	2.25	.12
			W	56	24.83	1.56	.06
	Oct	Minn	N	10	16.88	3.52	.21
			W	77	25.49	1.59	.06
		Ohio	N	28	18.49	2.88	.16
			W	55	22.95	1.66	.07
Plants per 6" section (average row 1 & 2)	Aug	Minn	N	7	1.50	.55	.37
			W	43	3.95	.27	.07
		Ohio	N	22	1.75	.21	.12
			W	26	4.35	.30	.07
	Sept	Minn	N	14	1.54	.27	.18
			W	81	3.72	.20	.05
		Ohio	N	36	1.81	.14	.08
			W	58	4.12	.23	.06
	Oct	Minn	N	13	1.58	.29	.18
			W	77	3.72	.20	.05
		Ohio	N	34	1.74	.17	.10
			W	57	3.98	.20	.05

1/ Counts were computed from the 6-inch section adjacent to the 3-foot section as in the operational soybean objective yield.

2/ N equals row width < 1.5 foot.
W equals row width ≥ 1.5 foot.

3/ CV = (Std. error of mean)/mean