United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research and
Applications
Division

NASS Staff Report
Number SRB-87-06

July 1987

Effect of Harvest Delays on Sunflower Yield Estimates

Ralph V. Matthews
Douglas C. Bond

EFFECT OF HARVEST DELAYS ON SUNFLOWER YIELD ESTIMATES by Ralph V. Matthews and Douglas C. Bond, Research and Applications Division and Estimates Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250, July 1987. Staff Report No. SRB-87-06.

ABSTRACT

The National Agricultural Statistics Service conducts sunflower objective yield surveys to forecast and estimate the yield and production of the oil-type sunflower crop. Survey procedures require the harvest of sample plots for yield estimation no more than 7 days before the farmer harvests the sample field. Experiments were conducted in 1984 and 1985 at North Dakota State University to determine the effect of a delay in harvest of 10, 20, or 30 days. No significant differences were detected among the harvest dates in either year due to large standard deviations of approximately 10 and 17 percent of the overall mean in 1984 and 1985, respectively. Covariance analyses with either stalk number or head number as the covariate were very similar in 1985. Differences between years in the average yield and the residual mean squares precluded a combined, 2-year analysis.

KEY WORDS: sunflower objective yield, yield estimation, analysis of variance, covariance analysis, power

ACKNOWLEDGMENTS

Al Schneiter, Professor of Agronomy, North Dakota State University, conducted the field experiments and supervised the lab work. Bob Battaglia, Fatu Bigsby, Ron Steele, and Ron Fecso made suggestions which clarified the presentation of the results.

CONTENTS

SUMMARY

The sunflower objective yield surveys of the National Agricultural Statistics Service are conducted to forecast and estimate oil-type sunflower yield and production. Enumerators are instructed to harvest sample plots for yield estimation no more than 7 days before farmer harvest. The objective of this research was to determine if sunflower yield estimates changed in response to harvest delays of 10, 20, and 30 days.

In 1984 and 1985, experiments were conducted at North Dakota State University to compare yield estimates from plots harvested 10, 20, and 30 days after the first harvest. The 4 harvest-day treatments were compared in an experiment with a randomized complete block design having 16 replications. No significant differences occurred among harvest dates in either year for a model with or without the number of stalks as a covariate. The stalk number covariate was significant (alpha = .05) in 1985. Nonsignificant interactions between replications and the covariate indicated a constant yield response to the covariate in all replications in each year. In 1985, covariance models with stalk number or head number covariates had nearly identical results and regression diagnostics.

Relatively large coefficients of variation (approximately .10 in 1984 and .17 in 1985) meant that the experiments could not detect small differences which are important to the sunflower objective yield program. Ambiguity remains about whether time lags between the enumerator's harvest and the farmer's harvest contribute to operationally important differences in yield estimates.

Data from the 2 years were not combined for a single analysis due to non-homogeneity of the residual mean squares. This difference between years was also shown by an average yield in 1985 which was 19 percent less than that of 1984.

EFFECT OF DELAYS IN HARVESTING ON SUNFLOWER YIELD ESTIMATES

By Ralph V. Matthews and Douglas C. Bond[1]

INTRODUCTION

The sunflower objective yield survey of the National Agricultural Statistics Service began on an experimental basis in North Dakota in 1981 [2].[2] The survey became operational and was expanded to Minnesota and South Dakota in 1984. A research report on forecasting and estimating improvements [1] recommended a study of the time lags between the enumerator's harvest and the farmer's harvest. It was hypothesized that a time gap between the two harvests caused the OY yield estimate to be a biased estimate of the farmer-harvested crop. Experiments were conducted for NASS in 1984 and 1985 by North Dakota State University researchers to determine if delays in harvest affected the yield estimates.

BACKGROUND

The sunflower Enumerator's Manual [7] states that the enumerator's final preharvest visit should be made within 7 days of the farmer's harvest. The date of the farmer's harvest is collected in a post-harvest personal or telephone interview during which the farmer-reported yield is also collected. Results of the 1983 OY survey in North Dakota showed that 81 percent of the sample units were visited more than 7 days before the farmer's harvest (table 1). Twenty-six percent of the sample units were harvested more than 21 days before the farmer-reported harvest date.

Table 2 contains the harvest delay results of the 1984 and 1985 OY surveys. In 1984, 77 percent of the sample units were harvested by the enumerator more than 7 days before the farmer's harvest; in 1985 the figure was 56 percent. The percentages of

[1] The authors are mathematical statisticians with the National Agricultural Statistics Service, U.S. Department of Agriculture.

[2] Numbers in brackets refer to literature cited at the end of this report.

sample units harvested by the enumerator more than 21 days before the farmer's harvest were 30 percent and 7 percent in 1984 and 1985, respectively.

Table 1 -- Number of days the enumerator's harvest date preceded the farmer-reported harvest date; 1983 sunflower OY survey, North Dakota

Number of days	Number of samples	Percent of total
0 to 7	21	19
8 to 14	46	43
15 to 21	13	12
22 to 28	14	13
29 to 35	8	7
Over 35	6	6
TOTAL	108	100

Table 2 -- Number of days the enumerator's harvest date preceded the farmer-reported harvest date; 1984 and 1985 sunflower OY surveys, North Dakota, Minnesota, and South Dakota

	1984		1985	
Number of days	Number of samples	Percent of total	Number of samples	Percent of total
0 to 7	36	23	60	44
8 to 14	36	23	41	30
15 to 21	37	24	24	18
22 to 28	26	17	8	6
29 to 35	13	8	2	1
Over 35	8	5	0	0
TOTAL	156	100	135	99[1]

[1] Total does not equal 100 due to rounding.

Delays of this magnitude could produce overestimates of the farmer's yield if results in sunflower are similar to those for corn [8]. Losses of 5 to 15 percent of final dry matter were observed when the harvest in hybrid corn test plots was delayed by 24 days. The losses appeared related to continued respiration and the conversion of stored seed reserves during field drying versus a quick cessation of respiration with artificial drying. The experiments described herein examined the possibility of a similar situation with sunflower.

METHODOLOGY

The experiments were conducted in 1984 and 1985 near Fargo, North Dakota. Hybrid 894 was planted in 30-inch rows at a higher than recommended density and thinned by hand to 20,000 plants per acre at the V-2 stage [5] of development to ensure equal plant populations. Each experiment was a randomized complete block design with 16 replications. Each replication consisted of four 15-foot sections of one row which were assigned randomly for harvest at one of four harvest dates. To prevent the harvest of one plot from affecting the remaining plots, border rows were left intact between plots. The first harvest occurred when the crop was mature by the objective yield survey definition: plants almost totally brown and brittle and ready for harvest within 7 days [7]. The other harvest dates were 10, 20, and 30 days later.

The number of plants in each plot was counted in both years, and the number of heads was also counted in 1985. The heads were harvested by hand and dried in a forced-air dryer before threshing. The seeds from each plot were weighed after blank seeds and trash were removed. In 1985, two subsamples of approximately 20 g were weighed, oven dried, and reweighed to determine the moisture content after the forced-air drying.

The weight of filled seeds per plot (g) was converted to the gross yield (lb/ac) with the following formula:

$$\text{gross yield} = \frac{\text{seed weight} * 43560}{2.5 * 15 * 453.6}$$

where seed weight = filled seeds per plot (g)
43560 = ft^2 per acre
2.5 = rowspace (ft)
15 = plot length (ft)
453.6 = g per lb.

The gross yield variable was analyzed by analysis of variance and analysis of covariance to evaluate the effect of delays in harvesting.

The analysis of variance model was

$$Y_{ij} = U + R_i + D_j + E_{ij}$$

where
Y_{ij} = Gross yield
U = Overall mean
R_i = Replication (1 to 16)
D_j = Day (0,10,20,30)
E_{ij} = Error term.

The assumptions of this model are that the replications and days are independent and that error variance is constant among replications.

The covariance model was

$$Y_{ij} = U + R_i + B(X_{ij}) + R_i{*}B(X_{ij}) + D_j + E_{ij}$$

where
Y_{ij} = Gross yield
U = Overall mean
R_i = Replication (1 to 16)
B = Regression coefficient between Y and X
X_{ij} = Number of stalks per harvested plot
$R_i{*}B(X_{ij})$ = Interaction between replications and the stalk number covariate
D_j = Day (0,10,20,30)
E_{ij} = Error term.

The assumptions of this model are that the replications, days, and covariate are independent. The interaction term provides a test of the independence of replications and the covariate. Error variance among replications is also assumed constant. The assumption was met in 1984, but not in 1985 due to one replication with a much larger variance than all the other replications. One plot yield which was more than 2 standard deviations below the overall mean appeared responsible for this large variance. No evidence indicated that the low yield was an invalid value, and it was not removed from the analysis.

RESULTS

Table 3 contains the 1984 gross yield means and the 1985 gross yield and moisture content means. Each value is a mean of 16 observations. The experiment-wide means showed that the 1985 gross yield was 19 percent lower than that of 1984. The 1985 crop was affected by lower than normal temperatures during June, July, and August.

Table 3 -- Gross yield and moisture content means, 1984 and 1985 sunflower harvest delay experiments; Fargo, North Dakota

	1984		1985		
Days	Gross yield (lb/ac)	Change (%)	Gross yield (lb/ac)	Change (%)	Moisture content (%)
0	1,967	-	1,490	-	3.5
10	1,855	-5.7	1,589	+6.6	3.9
20	1,840	-.8	1,522	-4.2	3.3
30	1,820	-1.1	1,454	-4.5	3.6
All	1,870		1,514		3.6

In 1984, a decrease in the gross yield estimate occurred for each 10-day delay in harvesting. This trend was not present in 1985, since the 10- and 20-day harvests both yielded more than the initial harvest. The percentage changes from the previous harvest dates were not uniform for the 10-day increments.

The moisture data in 1985 showed that differences occurred in the four harvest dates, but the changes in moisture content did not cause corresponding changes in gross yield. Although variation was present in the moisture content values, they were much lower than the range of moisture content at harvest, which is 15 to 20 percent. Moisture content was not used to adjust to a standard moisture percent, because moisture content data was not available for 1984.

Table 4 shows the analysis of variance for the 1984 experiment. The sources of variation were the replications and the harvest day treatments. The yield estimates from the 4 harvest days were not statistically different. The coefficient of determination (R^2) and the coefficient of variation (CV) were included for comparisons with other models.

Table 4 -- Analysis of variance of gross yield, 1984 sunflower harvest delay experiment; Fargo, North Dakota

Source	df	Sum of squares	Mean square	F	Pr > F
Blocks	15	2,252,073.87	150,138.26	3.40	<.01
Days	3	206,833.97	68,944.66	1.56	.21
Residual	45	1,988,020.08	44,178.22		
Total	63	4,446,927.92			

R^2 = .55
CV = .11

Table 5 contains the 1984 covariance analysis with the number of stalks per harvested plot as a covariate. The partial or Type III sums of squares [4] were calculated as if each effect entered the model last. The covariate term and its interaction with replications were both nonsignificant. The covariate's lack of significance indicated that little was gained over the analysis in table 4. The nonsignificance of the interaction confirmed the assumption of independence between the replications and the covariate. The harvest day effect was nonsignificant, which confirmed the results in table 4.

Table 5 -- Covariance analysis of gross yield with stalk number as a covariate, 1984 sunflower harvest delay experiment; Fargo, North Dakota

Source	df	Sum of squares	Mean square	F	Pr > F
Blocks (B)	15	672,903.23	44,860.22	1.44	.20
Stalks (S)	1	59,705.10	59,705.10	1.91	.18
B * S	15	680,648.30	45,376.55	1.45	.19
Days	3	94,232.41	31,410.80	1.01	.40
Residual	29	904,795.40	31,199.84		

R^2 = .80
CV = .09

The covariance analysis was preferable because stalk number per plot, a factor which affects yield, was included. This analysis also had a higher R^2 value and a slightly lower CV. Both these changes resulted from the extra terms in the model which removed a portion of the residual sum of squares.

The 1985 analysis without the stalk number covariate is shown in table 6. The conclusion was that no significant difference occurred due to differences in harvest time. The 1984 (table 4) and 1985 (table 6) analyses differed in a number of factors. In 1985, the mean and R^2 values were lower, the CV was higher, and the block effect was nonsignificant. Despite these year to year differences, a nonsignificant difference due to harvest delay was observed in both years.

Table 6 -- Analysis of variance of gross yield, 1985 sunflower harvest delay experiment, Fargo, North Dakota

Source	df	Sum of squares	Mean square	F	Pr > F
Blocks	15	1,012,076.51	67,471.77	.95	.52
Days	3	157,360.10	52,453.37	.74	.54
Residual	45	3,201,844.32	71,152.10		
Total	63	4,371,280.93			

R^2 = .27
CV = .18

The two widely different years confirmed that testing in 2 years was important. Reaching the same conclusion in the 2 years reinforced the conclusion of no statistical difference among the harvest days.

Table 7 has the 1985 covariance analysis with the stalk number per harvested plot as a covariate. No significant difference due to days was observed. The covariate was significant, which indicated an improvement over the analysis in table 6. The utility of the stalk number covariate was thus shown in 1985 but not in 1984. The interaction between replications and the covariate was again nonsignificant.

To test the influence of a possible outlier, a least squares estimate [6 p.275] was calculated for that cell, and the data were re-analyzed. In that case, the test statistics for blocks and the blocks by stalks interaction were significant at the 10 percent alpha level. The significance of the stalk number covariate and the nonsignificance of the harvest day treatments were unchanged in the alternative analysis. Tables 6, 7, and 8 are based on the original, unaltered data.

Table 7 -- Covariance analysis of gross yield with stalk number as a covariate, 1985 sunflower harvest delay experiment, Fargo, North Dakota

Source	df	Sum of squares	Mean square	F	Pr > F
Blocks (B)	15	932,439.82	62,162.65	.97	.50
Stalks (S)	1	352,912.89	352,912.89	5.52	.03
B * S	15	973,697.50	64,913.17	1.02	.47
Days	3	20,706.63	6,902.21	.11	.95
Residual	29	1,852,891.16	63,892.80		

R^2 = .58
CV = .17

Since the number of heads harvested was also counted in 1985, another covariance analysis was possible with the number of heads per plot as a covariate (table 8). This analysis was similar to the analysis in table 7 with respect to the nonsignificance of the day effect, the significance of the covariate, the R^2 value, and the CV. The close similarity between stalk number and head number covariance analyses in 1985 indicated that stalk number was an acceptable covariate.

Table 8 -- Covariance analysis of gross yield with head number as a covariate, 1985 sunflower harvest delay experiment, Fargo, North Dakota

Source	df	Sum of squares	Mean square	F	Pr > F
Blocks (B)	15	842,092.14	56,139.48	.85	.62
Heads (H)	1	329,135.45	329,135.45	4.98	.03
B * H	15	866,908.92	57,793.93	.88	.60
Days	3	24,199.59	8,066.53	.12	.95
Residual	29	1,914,738.92	66,025.48		

R^2 = .56
CV = .17

Pearson-Hartley power charts [3] were used to determine the power of the stalk number covariance analyses for an alpha level of .05 and 16 replications. For three equal treatment means and a fourth which is one standard deviation greater, the probability is at least .79 of rejecting the null hypothesis of equal means. The square root of the residual mean square (1984, 177 lb/ac; 1985, 253 lb/ac) provided an estimate of the standard deviation in each year. The coefficients of variation were 9.5 percent and 16.7 percent in 1984 and 1985, respectively. A differences among treatments of 5 percent could not be detected due to high CV's, especially in 1985, and the conclusion of no significant differences due to delay in harvest must be made with caution.

In table 9, the residual mean squares of the 1984 and 1985 stalk number covariance analyses were compared using an F test. If the two years do not have homogeneous errors, they should not be analyzed together. The F test showed that the residual terms were not equal, and a combined analysis was not performed. Although this was further evidence that the two years differed, the conclusion of no difference among the four harvest dates was reached in each year.

Table 9 -- Comparison of the residual mean squares from stalk number covariance analyses of gross yield, 1984 and 1985 sunflower harvest delay experiments, Fargo, North Dakota

Year	df	Residual mean square	F	Pr > F
1985	29	63,892.80	2.05	.03
1984	29	31,199.84		

CONCLUSIONS

Gross yield estimates using sunflower objective yield procedures were not significantly different on harvest dates which were 10, 20, and 30 days after the first harvest. The yields for the 30-day delayed harvests were 7 and 2 percent less than the initial harvests in 1984 and 1985, respectively. In 1984, the yields decreased through the 30-day period. In 1985, the yields for the 10- and 20-day delayed harvests exceeded the initial harvest yield by 7 and 2 percent, respectively.

The conclusion of no difference was reached in both experiments, despite 1985 yields that averaged 19 percent less than 1984 yields. A combined, 2-year analysis was not possible due to non-homogeneity of the residual mean squares. Recent changes in operational procedures have been based on analysis of only one year's test results. This study indicated that year to year variation affects test results and testing of procedures should be done in more than one year.

A stalk number covariate was significant in 1985 (alpha = .05), and no significant interaction was detected between replications and the covariate in either year.

Due to high CV's, the experiments could not detect small differences which are important to the sunflower objective yield program. It remains unclear whether time lags between the enumerator's harvest and the farmer's harvest contribute to operationally important differences in yield estimates.

RECOMMENDATIONS

1. Make no changes regarding the number of days between enumerator harvest and farmer harvest. If the 7-day limit is exceeded, there is little effect on the yield estimates.

2. Test hypotheses about procedures in more than one year to ensure the validity of the conclusions.

3. Do not extrapolate the results of this experiment to other objective yield crops without evidence that changes do not occur with delays in harvest.

REFERENCES

1. Bond, Douglas C. 1983 Sunflower Objective Yield Survey Research. U.S. Dept. of Agriculture, Statistical Reporting Service. June 1985.

2. Craig, Michael E. 1981-82 Sunflower Objective Yield Research. U.S. Dept. of Agriculture, Statistical Reporting Service. May 1984.

3. Guenther, William C. Analysis of Variance. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964.

4. SAS Institute, Inc. SAS(R) User's Guide: Statistics, Version 5 Edition. Cary, North Carolina: SAS Institute, Inc., 1985.

5. Schneiter, A. A., and J. F. Miller. "Description of Sunflower Growth Stages." Crop Sci., 21:901-903. 1981.

6. Snedecor, George W., and William G. Cochran. Statistical Methods, Seventh ed. Ames, Iowa: Iowa State University Press, 1980.

7. Statistical Reporting Service. Enumerator's Manual, 1984 Sunflower Objective Yield Survey. U.S. Dept. of Agriculture, Statistical Reporting Service. June 1984.

8. Vogel, John. "Vanishing Yield Syndrome - Letting Corn Field Dry May Cost 5 to 20 Bushels." Prairie Farmer. Sept. 1, 1984.