COTTON YIELD AND GINNING STUDY

Gary N. Bovard

Statistical Research Division

Economics and Statistics Service

U. S. Department of Agriculture

November 1980

COTTON YIELD AND GINNING STUDY. By Gary N. Bovard; Economics and Statistics Service; U.S. Department of Agriculture

ABSTRACT

This paper examines three components of the Cotton Objective Yield Survey -- the seed lint ratio, the harvested seed cotton estimate and the harvest loss estimate. Possible causes for differences between objective yield estimates of seed cotton and gin reported seed cotton yields may be harvest loss located in gin trash from stripper harvested fields and the time lag between the final preharvest and the postharvest field measurements. The relationship between lint yield and seed cotton yield can be represented by a ratio. Harvest loss directly outside field boundaries may approach 1.0 percent of biological yield in stripper harvested fields.

Keywords: cotton, seed lint ratio, gin trash.

ACKNOWLEDGEMENTS

Numerous present and past members of the research staff have contributed to this study, but special recognition should go to Bill Iwig, Ben Klugh and Jack Nealon for their study design and data collection efforts.

CONTENTS

Cotton Yield and Ginning Study

Gary N. Bovard

INTRODUCTION

The Cotton Objective Yield Survey, as conducted by the Statistics Unit of the Economics and Statistics Service (ESS), USDA produces estimates of harvested lint cotton yield per acre at state, regional and national levels. During the past few years, the final objective yield estimates have often exceeded the Crop Reporting Board estimated yields at all levels. The Cotton Yield and Ginning Study was conducted in 1976 and 1977 to examine current Cotton Objective Yield estimating procedures.

This study was designed to examine three components of the Cotton Objective Yield Study -- the seed lint ratio, the harvested seed cotton estimate and the harvest loss estimate. The seed lint ratio is a parameter used in the objective yield program to convert objective estimates of seed cotton yield to lint yield. This paper examines the appropriateness of using a ratio to represent the relationship between seed cotton yield and lint yield. The harvested seed cotton yield estimate is the difference between the biological yield estimate and harvest loss estimate. The harvested seed cotton estimate is compared to a gin reported seed cotton yield in an attempt to quantify the size of the nonsampling error and to identify possible causes of the nonsampling error. Objective yield harvest loss estimates are based only on harvest loss located within the field boundaries. The study also measures the amount of harvest loss located directly outside the field boundaries.

SAMPLE DESIGN

The Cotton Yield and Ginning Research Study was conducted in the West Texas Panhandle near Lubbock and in the Mississippi Delta near Greenville. These study areas are referred to as "Lubbock" and "Greenville" in this paper. The two areas were chosen to represent two different harvesting methods. At Lubbock, cotton fields are harvested once by cotton strippers. At Greenville, where conventional pickers are used, many fields may be harvested two or more times. The study was conducted over two years to provide varying growing and harvesting conditions for the analysis. In both years, 25 fields were observed in each region. A sample of five fields was drawn from each of five selected gins in each area.

Objective measurements of biological yield components were taken from 20 preharvest plots in each field in 1976 and from 30 preharvest plots in 1977. Only one preharvest visit was made prior to each harvest. objective yield survey procedures were applied in the 10-foot by 2-row preharvest plots. Postharvest observations were made within a few days following harvest. Within each field, 10 postharvest plots were located 2 rows to the right or left of 10 randomly selected preharvest plots. Postharvest plots were also randomly located outside the ends and sides of each field. Five plots were located at the ends of each field in 1976. In 1977, this number was reduced to 4. Two postharvest plots were located along the sides of each field both years.

At Lubbock, a two-stage sampling procedure was used to estimate harvest loss in dump piles. The initial stage involved randomly sampling a number of dump piles based on the total number of dump piles within and outside the field boundaries. Dump piles were defined as boll concentrations meeting either of two criteria:

1. the boll concentration covered an area of at least 16 square feet, or

2. bolls were piled to a depth of six inches or more and the concentration covered an area of at least four square feet.

All material in a sampled dump pile was then gathered into buckets. Several buckets were randomly sampled based on the total number of buckets gathered from a pile. Harvest loss material in each bucket was then measured. Additional information on dump pile sampling can be found in the enumerator's manual for this study.

Operators reported the harvested cotton acreage for each sample field. The cotton from the sampled field was harvested and ginned separate

from all other fields of the operator. Bale weights of ginned lint cotton from each field were recorded at the gin. Cotton seed weights were obtained at Lubbock both years and at Greenville in 1976. At Greenville in 1977, total seed cotton weight was calculated from the lint weight using a gin determined factor, the inverse of each gin's seed lint ratio.

Of the 25 fields sampled each year, one field at Lubbock in 1976 was not harvested and gin data for four fields in the 1977 Greenville data set were incomplete. These fields were excluded fron the analysis.

DATA ANALYSIS

Seed Lint Ratio

The Cotton Objective Yield Survey procedures convert seed cotton yield estimates to lint yield using a seed lint ratio. The seed lint ratio in the model is a three-year average of seed lint ratios calculated annually from gin reported ginnings data. The use of the seed lint ratio as a parameter in the model assumes that the ratio of lint yield to seed cotton yield is relatively constant between recent years and is not related to yield level.

In this study, both lint and seed production were measured by the gins for each field, except, at Greenville in 1977 where seed cotton production was based on a gin determined ratio and the weight of the lint cotton. Simple linear regression models regressing lint yield on seed cotton yield are presented in Table 1.

Intercepts which are not significantly different from zero, and high R^2 values, imply that a ratio can be used to represent the actual relationship between lint yield and seed cotton yield within year. The results in Table 1 should not be used to conclude that the seed lint ratio used in the objective yield model is correct since the nonrandom method of selecting fields excludes a between-year comparison of ratios.

Table 1: Simple Linear Regression Equation, Lint Yield Versus Gin Measured Seed Cotton Yield, Pounds per Acre

Area	Year	Equation 1/	R^2
Lubbock	1977	Lint = - 7.9 + .3857 Seed Cotton	.995
Lubbock	1976	Lint = - 0.5 + .3676 Seed Cotton	.992
Greenville	1976	Lint = 11.4 + .3793 Seed Cotton	.996

1/ No intercepts were significantly different from zero at the 95% level of probability.

Seed Cotton Estimates

Data collection procedures and equations used to form field level seed cotton estimates in this study were very similar to the Cotton Objective Yield procedures and equations. As such, inferences can be made about the objective yield survey based upon the results of this study. However, the study was not designed as a validation study of the Cotton Objective Yield Survey. Listed below are some of the field observation differences which exist between the objective yield procedures and those used in this study:

1. Only one preharvest visit was made in most fields of the study, whereas the objective yield survey has several monthly visits. Any bias in the objective yield data set due to previous handling of plants would not be included in the study estimates.

2. In this study, harvest loss was measured following each harvest in multiple-harvest fields. In an objective yield survey, harvest loss is measured only after the final harvest.

3. Procedures used in this study included the sampling of harvest loss located outside field boundaries. The objective yield survey only measures harvest loss within the field boundaries.

Equations used to form biological seed cotton yield and harvest loss estimates are presented in the Appendix.

Field level gin reported seed cotton yield and objective yield estimates can be found in Tables 2 through 5. The objective yield estimates of seed cotton yield differed from gin reported seed cotton yield by more than 2 standard errors in 11 of the 25 Lubbock 1977 fields and 9 of the 24 Lubbock 1976 fields. At Greenville, 9 of the 21 estimates fell outside the 2 standard error confidence interval in 1977 and 8 of the 25 in 1976. At Lubbock, the tendency was toward overestimation with 17 of the 25 estimates in 1977, and 17 of the 24 estimates in 1976, being overestimates of the gin reported yield. At Greenville, only 8 of the estimates were too high each year.

Two nonsampling error sources were identified which may explain some of these differences. The time lag between the preharvest and postharvest visits showed some correlation with the differences. At Greenville, in 1977, the correlation (+0.58) between the time lag and the differences was significant at $\alpha = .05$. At Lubbock in 1977,

Table 2: Gin Reported Seed Cotton and Seed Cotton Estimates, Lubbock 1977

Field	Time Lag 1/	Gin Seed Cotton	Seed Cotton Estimate	Seed Cotton Difference 2/	Seed Cotton Standard Error	Modified Seed Cotton Estimate 3/	Gin Lint Yield
	- -days- -		- - - - - - - - - - pounds per acre- - - - - - - - - - - - - - - - - -				
1	30	2,849	3,330	481*	73	3,169	1,108
2	43	931	981	50	55	928	326
3	14	1,617	1,940	323*	80	1,839	614
4	7	1,854	1,897	43	106	1,768	680
5	15	1,640	1,840	200*	80	1,746	585
6	23	1,169	970	-199*	62	842	416
7	23	2,166	1,707	-459*	70	1,545	810
8	1	924	932	8	64	829	358
9	1	376	495	119*	39	506	137
10	3	301	277	- 24	32	285	98
11	22	835	881	46	50	870	325
12	19	925	921	- 4	50	879	351
13	8	1,384	1,655	271*	85	1,620	546
14	40	812	1,032	220*	65	967	319
15	14	2,505	2,923	418*	106	2,868	954
16	14	1,956	1,890	- 66	92	1,749	759
17	8	1,559	1,573	14	79	1,536	620
18	24	800	912	112	69	835	310
19	1	622	612	- 10	67	610	222
20	27	1,073	1,657	584*	64	1,625	418
21	7	1,232	1,428	196*	86	1,352	492
22	15	1,860	1,798	- 62	98	1,619	715
23	5	1,290	1,335	45	59	1,166	500
24	10	1,423	1,515	92	107	1,322	536
25	2	796	746	- 50	47	734	295
Mean	15	1,315	1,410	95	142	1,328	500

* Seed cotton estimate is more than 2 standard errors from the gin reported seed cotton.

1/ Days between preharvest and postharvest observations.

2/ Difference = seed cotton harvested yield estimate - gin reported seed cotton yield.

3/ Large unopen and partially open bolls excluded from the estimate.

Table 3: Gin Reported Seed Cotton and Seed Cotton Estimates, Lubbock 1976

Field	Time Lag 1/	Gin Seed Cotton	Seed Cotton			Modified Seed Cotton Estimate 3/	Gin Lint Yield
			Estimate	Difference 2/	Standard Error		
	days	pounds per acre					
1	23	921	934	13	99	598	344
3	2	1,392	1,378	- 14	127	1,122	548
4	10	1,128	1,428	300*	116	1,471	451
5	1	1,956	2,248	292*	88	2,090	708
6	1	1,139	1,165	26	92	1,005	407
7	6	828	764	- 64	94	686	307
8	4	360	378	18	46	271	121
9	5	661	677	16	71	532	235
10	11	109	22	- 87*	17	31	38
11	42	2,212	2,240	28	100	1,740	794
12	28	1,496	1,574	78	106	1,312	536
13	21	1,447	1,413	- 34	116	1,182	532
14	19	1,439	1,374	- 65	124	1,235	525
15	4	1,585	1,801	216*	102	1,515	569
16	25	437	501	64	52	366	159
17	8	1,471	1,726	255*	58	1,471	572
18	2	1,378	1,480	102	93	1,109	531
19	26	915	1,091	176*	85	846	313
20	14	1,016	1,413	397*	116	1,193	354
21	23	1,269	1,014	-255*	110	1,000	469
22	13	1,309	1,432	123	127	1,349	460
23	12	2,088	1,492	-596*	113	1,390	766
24	1	1,031	1,085	54	84	968	385
25	3	1,118	1,195	77	80	1,091	416
Mean	13	1,196	1,242	46	108	1,066	439

* Seed cotton estimate is more than 2 standard errors from the gin reported seed cotton.

1/ Days between preharvest and postharvest observations.

2/ Difference = seed cotton harvested yield estimate - gin reported seed cotton yield.

3/ Large unopen and partially open bolls excluded from the estimate.

Table 4: Gin Reported Seed Cotton and Seed Cotton Estimates, Greenville 1977

Yield	Time Lag 1/	Gin Seed Cotton	Seed Cotton			Modified Seed Cotton Estimate 3/	Gin Lint Yield
			Estimate	Difference 2/	Standard Error		
	days		pounds per acre				
1	14	2,017	2,137	120	112	1,824	780
2	10	1,723	1,651	72	99	1,488	666
3	14	1,384	1,229	-155*	62	1,065	535
4	1	1,731	1,494	-237	135	1,627	667
5	5	1,756	1,295	-461*	143	1,151	679
8	1	1,147	918	-229*	95	873	457
9	8	1,674	1,513	-161*	64	1,444	645
10	6	900	816	- 84	72	769	345
11	13	821	587	-234*	112	586	316
12	9	755	612	-143*	64	606	290
13	23	1,160	1,267	107	60	1,149	446
16	15	2,014	2,125	111	162	1,833	777
17	23	1,289	1,589	300*	82	1,205	496
18	16	2,409	3,034	625*	128	2,496	929
19	8	2,144	1,933	-211	216	1,685	826
20	15	1,934	1,939	5	158	1,375	746
21	19	1,331	1,636	305*	138	1,511	512
22	5	1,563	1,506	- 57	104	1,549	599
23	8	3,167	3,435	268	262	2,909	1,216
24	22	2,400	2,355	- 45	229	1,744	923
25	3	2,150	1,874	-276	184	1,702	828
Mean	11	1,689	1,664	- 25	156	1,457	651

* Seed cotton estimate is more than 2 standard errors from the gin reported seed cotton.

1/ Days between preharvest and postharvest observations. First harvest.

2/ Difference = seed cotton harvested yield estimate - gin reported seed cotton yield.

3/ Large unopen and partially open bolls excluded from estimate.

Table 5: Gin Reported Seed Cotton and Seed Cotton Estimates, Greenville 1976

Field	Time Lag 1/	Gin Seed Cotton	Seed Cotton			Modified Seed Cotton Estimate 3/	Gin Lint Yield
			Estimate	Difference 2/	Standard Error		
	-days		-pounds per acre-				
1	13	369	360	- 9	59	260	143
2	19	118	50	- 68	44	34	43
3	24	67	20	- 47*	18	16	25
4	0	391	288	-103	57	317	142
5	11	680	389	-291*	106	303	270
6	12	1,511	1,473	- 38	115	1,038	561
7	7	2,383	2,734	351	179	2,128	925
8	26	2,440	2,594	154	197	2,265	916
9	13	1,920	1,794	-126	285	1,470	715
10	4	1,572	1,798	226*	88	1,763	589
11	12	1,535	1,822	287	156	1,342	596
12	17	441	357	- 84	73	228	167
13	2	768	720	- 48	137	687	292
14	4	1,221	1,136	- 85	112	1,068	494
15	19	1,620	2,214	594*	200	1,296	630
16	7	104	45	- 59*	13	41	42
17	13	542	520	- 22	67	472	239
18	2	473	418	- 55	50	410	199
19	10	498	579	81	51	535	216
20	2	590	637	47	72	616	251
21	18	894	755	-139	140	494	358
22	4	1,800	1,187	-613*	100	1,168	731
23	4	616	301	-315*	78	226	246
24	10	1,065	1,058	- 7	91	754	430
25	1	949	1,434	485*	147	1,454	383
Mean	10	982	987	5	160	815	384

* Seed cotton estimate is more than 2 standard errors from the gin reported seed cotton.

1/ Days between preharvest and postharvest observations. First harvest.

2/ Difference = seed cotton harvested yield estimate - gin reported seed cotton yield.

3/ Large unopen and partially open bolls excluded from estimate.

the time lag was not significantly related to the differences but the correlation (+0.54) between the time lag and the absolute value of the differences was significant at $\alpha = .01$. There was little correlation between time lag and differences at either location in 1976. The lack of a significant relationship in 1976 may be due to the short time lag experienced for a large number of fields. In 1976, there were eight fields at Lubbock and nine fields at Greenville for which the time lag was less than five days. In 1977, only five fields at Lubbock and three fields at Greenville had a time lag of less than five days.

A second factor which may cause overestimation is the possible inclusion of large unopen and partially open bolls in gin trash when the cotton has been harvested with stripper machines. At Lubbock, where cotton was stripper harvested, there is a large reduction in the number of large unopen and partially open bolls found between preharvest and postharvest counts. The large difference in counts would be expected when the time lag is lengthy regardless of harvesting method. However, when the time lag is only a few days, it is unlikely that a large percentage of the bolls would open. Table 6 contains a listing of fields where the time lag between preharvest and postharvest visits was four days or less. Picker harvested fields around Greenville had a low disappearance of large unopen and partially open bolls. If we assume that the disappearances observed at Greenville are representative of those expected, then the disappearance at Lubbock does reflect unmeasured harvest loss. If this assumption is true, harvest loss would be too high. Gin trash was not sampled either year, but a total weight of gin trash was obtained for each field in 1976. In each of the Lubbock fields with a small time lag the gin trash weight exceeded the weight estimated for the disappearing bolls.

The effect of the unmeasured harvest loss on the harvested seed cotton estimate could be quite large. The average disappearance from the 9 Greenville 1976 fields listed in Table 6 was 14 percent of the preharvest bolls. The average disappearance from the 8 Lubbock 1976 fields was 83 percent. If the Greenville disappearance is what should be expected with a short time lag, then there is an average unexplained disappearance of about 24,000 bolls per acre at Lubbock. If the unexplained disappearance was included in the Lubbock harvest loss, the effect would be to increase the harvest loss estimate, thus reducing the seed cotton estimate by 240 pounds per acre in those fields.

Despite the difficulties observed at the field level, mean seed cotton estimates over all fields were within 4.0 percent of the mean gin reported cotton yield in three of the four data sets. Only the

Lubbock 1977 mean seed cotton estimate differed from the gin yield by a significant amount -- 7.2 percent or 95 pounds. The number of large unopen and partially open bolls observed at preharvest in 1977 at Lubbock was relatively small, less than half those observed in the Lubbock 1976 data set. Thus, estimating difficulties cannot be fully attributed to the effect of large unopen and partially open bolls. In fact, 11 of the 25 Lubbock 1977 modified field estimates of seed cotton yield, which excluded large unopen and partially open bolls from the estimate, exceeded the gin measured seed cotton yield. The implication is that the objective yield procedures may not accurately estimate the yield of fully mature cotton.

Table 6: Preharvest and Postharvest Estimates of Large Unopen and Partially Open Bolls Per Acre for Preharvest and Postharvest Visits Occurring Within a Four-Day Period

Field Number by Location	Large Unopen and Partially Open Bolls Per Acre			
	Time Lag	Pre-harvest	Post-harvest 1/	Percent Reduction
	-days-	-bolls-		-pct-
Lubbock, 1977				
8	1	18254	6485	
9*	1	152	1498	
10*	3	196	1177	
19	1	3966	3777	
25	2	1537	407	
5 fields				45
Lubbock, 1976				
3	2	33258	3044	
5	1	30571	7122	
6	1	34277	11467	
8	4	27775	6291	
15	4	26954	2647	
18	2	66522	4549	
24	1	31291	8182	
25	3	23756	3562	
8 fields				83
Greenville, 1977 2/				
4	1	63896	62156	
8	1	5451	2043	
25	3	84025	79979	
3 fields				6
Greenville, 1976 2/				
4	0	18284	17607	
10	4	26920	22882	
13	2	9941	7952	
14	4	33737	27783	
18	2	11628	10260	
20	2	17594	14211	
22	4	78430	70246	
23	4	10145	2704	
25	1	49348	46644	
9 fields				14

* Very low counts in these fields resulted in postharvest estimates being higher than preharvest.

1/ Postharvest estimates at Lubbock include dump pile estimates.

2/ First harvest only.

HARVEST LOSS

Field level estimates of harvest loss consisted of the loss found in four locations:

1. ten plots inside the field boundaries,
2. five plots in 1976 and four plots in 1977 outside the ends of each field,
3. two plots outside the sides of each field, and
4. dump piles, both inside and outside the field boundaries (Lubbock only).

Procedures used on the ten plots inside the field boundaries were essentially the same as those used in the operating program, except that the plots were not allowed to include dump piles. Plots along the ends and sides of a field provided estimates of harvest loss which might fall outside the field boundaries, and would not be sampled using objective yield procedures. Dump piles were sampled to identify harvest loss which may not be included in regular objective yield estimates. Harvest loss statistics are summarized in Table 7.

Very little harvest loss was located outside the field boundaries when excluding dump piles. The mean harvest loss around the field boundaries was less than 0.3 percent of biological yield in all locations and years. The amount of harvest loss around the outside of the field tended to be higher at Lubbock, probably due to spreading the trash collected when stripping the cotton.

Two assumptions led to special sampling of dump piles at Lubbock. The first assumption was that harvest loss in dump piles located outside field boundaries were not represented by objective yield estimating procedures. The second assumption was that some enumerators may occasionally adjust the location of field plots to avoid including dump piles in the plot. The first assumption is true since no harvest loss accumulations outside field boundaries are sampled using objective yield procedures. The second assumption cannot be tested, but may occur on occasion. Dump pile statistics are presented in Tables 8a and 8b. Over all fields, about one-fifth of the total harvest loss was located in dump piles. If both of the above assumptions were true, and no dump piles were sampled, the estimated seed cotton yield in 1977 at Lubbock would have been 2.2 percent higher.

The location of dump piles, inside or outside field boundaries, was recorded during the 1977 study at Lubbock, Texas. If only the first assumption is true, then only the harvest loss located in dump piles outside the field boundaries would not be accounted for by the objective yield procedures. Fourteen fields in 1977 had dump piles outside the field boundaries. The mean loss over these fields from dump piles outside the field was 21.9 pounds per acre -- 14.3 percent of total harvest loss, and 1.4 percent of biological yield in those fields. Over all 25 fields in 1977, the mean loss from dump piles outside the field was 12.3 pounds, 0.8 percent of biological yield.

Table 7: Harvest Loss Data

Source	Mean	Minimum	Maximum	Standard Error	Percent of Total Harvest Loss	Percent of Biological Yield
	-pounds per acre-				-percent	
			LUBBOCK 1977			
In Field	76.9	15.4	318.4	13.1	78.6	5.1
Field End	2.0	.3	7.8	.3	2.0	.1
Field Side	.6	0	2.2	.1	.6	.04
Dump Piles	18.4	0	68.0	4.0	18.8	1.2
Total Harvest Loss	97.9	15.4	329.2	13.6		6.5
			LUBBOCK 1976			
In Field	89.7	24.0	164.1	7.7	74.8	6.6
Field End	2.5	0	12.2	.6	2.1	.2
Field Side	.5	0	6.6	.1	.4	.04
Dump Piles	27.3	0	106.7	6.6	22.7	2.0
Total Harvest Loss	119.9	25.3	208.2	10.6		8.8
			GREENVILLE 1977[1/]			
In Field	336.9	76.5	900.1	55.5	99.6	27.3
Field End	.6	0	1.8	.1	.2	.05
Field Side	.6	0	3.4	.2	.2	.05
Total Harvest Loss	338.1	77.4	900.2	55.6		27.4
			GREENVILLE 1976[1/]			
In Field	251.7	44.2	794.4	41.5	99.6	31.5
Field End	.8	0	2.1	.12	.3	.1
Field Side	.1	0	1.3	.06	.03	.01
Total Harvest Loss	252.6	44.3	795.1	41.6		31.6

1/ Data from final postharvest visit only.

Table 8a: Dump Pile Harvest Loss Statistics, Lubbock 1976

Field	Total Piles	Number of Piles Sampled	Mean Harvest Loss Per Acre	Percent of Total Harvest Loss	Standard Error of Estimate	RANGE	
						Minimum Harvest Loss 1/	Maximum Harvest Loss 1/
			-pounds -	-pct-	- - - - - -	-pounds per acre-	- - - - - -
1	0	0	0	0	0	0	0
3	37	8	55.4	29	1.9	47.3	64.6
4	24	8	46.4	33	6.0	15.3	69.0
5	22	8	52.0	30	2.4	40.0	60.5
6	70	10	45.7	25	6.1	22.1	88.0
7	0	0	0	0	0	0	0
8	4	2	4.5	5	.1	4.5	4.6
9	7	3	6.2	10	.8	4.8	7.6
10	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0
13	12	5	5.7	3	.8	4.1	8.3
14	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0
16	11	5	9.0	24	1.6	3.8	12.4
17	37	8	79.3	65	10.1	22.5	118.1
18	32	8	59.0	41	8.6	28.3	102.6
19	0	0	0	0	0	0	0
20	44	9	14.3	33	2.5	7.0	27.0
21	32	8	9.3	11	2.4	3.0	20.7
22	82	10	70.6	34	19.9	29.2	246.2
23	2	2	22.0	17	4.9	17.1	27.0
24	46	9	69.4	43	9.2	41.8	132.9
25	30	8	105.5	56	9.8	72.7	162.0

1/ Per acre estimates for individual dump piles.

Table 8b: Dump Pile Harvest Loss Statistics, Lubbock 1977

Field	Total Piles	Number Piles Sampled			Harvest Loss/ Acre	Percentage of Total Harv Loss	Standard Error of Estimate	RANGE	
		Total	In Field	Out of Field				Minimum Harv Loss1/	Maximum Harv Loss1/
					- lbs -	- pct -	- - - - - pounds per acre - - - - -		
1	8	5	0	5	14.1	14	.9	10.9	16.3
2	6	5	0	5	5.8	7	1.1	3.2	9.2
3	32	15	0	15	15.5	14	2.0	5.2	33.3
4	19	12	0	12	21.3	21	4.6	8.0	56.5
5	89	35	17	18	67.8	33	10.3	10.1	326.6
6	41	20	20	0	5.3	2	.5	2.4	10.7
7	29	15	15	0	4.7	7	.6	2.9	12.4
8	85	35	32	3	47.7	35	3.8	11.9	100.8
9	40	15	15	0	9.4	26	1.5	3.5	26.4
10	0	0	0	0	0	0	0	0	0
11	33	15	2	13	8.7	14	1.3	2.1	18.6
12	12	8	0	8	9.5	28	1.0	5.1	12.5
13	8	5	0	5	7.5	4	1.3	5.3	12.0
14	2	2	0	2	2.4	5	1.6	.8	4.0
15	25	15	0	15	61.8	50	8.4	10.0	144.1
16	33	15	15	0	27.6	34	2.0	14.0	42.1
17	33	15	15	0	36.0	43	2.8	18.0	52.5
18	26	15	15	0	11.7	18	2.5	5.2	42.0
19	10	5	3	2	3.9	3	1.3	.8	8.0
20	0	0	0	0	0	0	0	0	0
21	50	20	20	0	56.4	62	2.3	42.3	90.5
22	2	2	2	0	2.1	4	.4	1.7	2.5
23	6	5	5	0	17.1	32	3.2	9.4	26.0
24	7	5	0	5	18.5	24	2.3	12.5	24.4
25	9	5	0	5	4.6	23	.4	3.5	6.2

1/ Per acre estimates for individual dump piles.

Evaluation of Modified Estimates

One of the concerns in the Cotton Objective Yield program is whether the current procedures correctly adjust for the effect of large unopen and partially open bolls. It is not known at the final preharvest visit whether these bolls will produce cotton at harvest because this is heavily dependent upon the number of days until actual harvest occurs and weather conditions. In addition, if the time until harvest is lengthy, additional fruit may be produced and be present during the postharvest visit. These bolls would not have been included in the final preharvest biological yield estimate.

If the estimates of seed cotton yield excluded counts of all large unopen and partially open bolls, the resulting procedure might provide a more consistent relationship between lint yield and estimated seed cotton yield. The correlation coefficients for the relationship between actual seed cotton yield and the modified and objective yield estimates are presented in Table 9.

The lower correlation coefficients for the modified procedure in all locations and years indicate that the modified procedure, excluding large unopen and partially open bolls from final preharvest counts, does not improve the consistency of the field level estimates. Since the objective yield procedure appears to do at least as well as the modification and has the advantage of including all available data, no further analysis appears warranted.

Table 9: Pearson Correlation Coefficients for Gin Measured Seed Cotton Versus Estimates of Seed Cotton

Area	Year	CORRELATION COEFFICIENTS	
		Modified Procedure	Objective Yield Procedure
Lubbock	1977	.943	.959
Lubbock	1976	.911	.922
Greenville	1977	.927	.943
Greenville	1976	.938	.949

SUMMARY

Objective field level estimates of seed cotton yield obtained in this study varied from the gin reported yields by more than two standard errors in many fields. However, the average estimate over all fields was quite good in three of the four data sets. Three possible sources of field level error were identified which may cause the differences between the estimated and actual yields. First, the length of time between the final preharvest measurement and the postharvest measurement shows some positive correlation with the differences. Second, harvest loss may be underestimated for fields which are stripper harvested. The harvest loss count of large unopen and partially open bolls was lower than expected in many fields. It is suspected that harvest loss attributed to these fruit forms are being delivered to the gin as part of trash. Third, some harvest loss in stripper harvested fields is not accounted for when dump piles are located outside the field boundaries. At Lubbock in 1977, the total harvest loss located directly outside field boundaries amounted to about 15 pounds per acre, or nearly 1.0 percent of biological yield.

The regression of lint yield on gin measured seed cotton yield showed that a ratio could accurately represent the relationship between lint yield and seed cotton estimates if accurate, unbiased estimates of seed cotton yield were produced by objective yield procedures. However, the seed lint ratio used in the Cotton Objective Yield Program was not evaluated in this study.

RECOMMENDATIONS

Three problems were identified in this study which may warrant future study -- harvest loss located in trash at gins, harvest loss located in dump piles outside field boundaries, and the time lag between preharvest and postharvest observations. The time lag problem is the only problem affecting both the Greenville and Lubbock areas and may be resolved by minimizing the time lag in the operating program. The harvest loss problems do need to be studied further, but it may be advisable to review the entire Cotton Objective Yield Survey procedure before investing resources in further research.

The field level harvested yield estimate in pounds per acre is:

$$Y_i = \sum_{k=1}^{n_k} (\bar{X}_{1ik.} - \sum_{t=2}^{4} \bar{X}_{tik.} - \bar{X}_{5_{ik..}})$$

where

i = field number,

k = harvest number,

n_k = total number of harvests in field i, n_k = 1 or 2,

Y_i = harvested yield estimate for field i,

$\bar{X}_{1ik.}$ = average biological yield estimate for field i, harvest k,

$\bar{X}_{2ik.}$ = average in-field harvest loss,

$\bar{X}_{3ik.}$ = average end-of-field harvest loss,

$\bar{X}_{4ik.}$ = average side-of-field harvest loss,

$\bar{X}_{5ik.}$ = dump pile harvest loss per acre (harvest loss = 0 at Greenville).

Biological Yield Estimates

Each field level estimate of biological yield for each harvest is the mean of the plot level estimates of yield per acre. The biological yield equation is:

$$\bar{X}_{1ik.} = \frac{\sum_{b=1}^{n_b} X_{1ikb}}{n_b}$$

where

b = plot number

n_b = number of biological yield plots in a field (20 in 1976, 30 in 1977),

X_{likb} = estimate of biological yield per acre in plot b.

The plot level biological yield equation is:

$$X_{likb} = B_{ikb} W_{ikb} E_{ib}$$

where

B_{ikb} = number of burrs and open bolls in plot + number of large unopen and partially open bolls in plot,

E_{ib} = expansion factor converting estimate to pounds per acre,

W_{ikb} = mean open boll weight in plot adjusted to 5% moisture and adjusted for number of large unopen and partially open bolls.

$$W_{ikb} = \begin{cases} \frac{C_{ikb}}{O_{ikb}} A_{irb} & \text{if } \frac{O_{ikb}}{B_{ikb}} > .85 \\ \frac{C_{ikb}}{O_{ikb}} A_{ikb} \left(a_o + a_1 \frac{O_{ikb}}{B_{ikb}} \right) & \text{if } \frac{O_{ikb}}{B_{ikb}} \leq .85 \end{cases}$$

where

C_{ikb} = total weight of cotton from open bolls

O_{ikb} = number of open bolls

A_{ikb} = factor converting weight to 5% moisture

a_o and a_1 = coefficients from historic objective yield data

Harvest Loss Estimates

Field level estimates of in-field (t=2), end-of-field (t=3), and side-of-field (t=4) harvest loss per acre are means of the plot level estimates of harvest loss per acre.

$\bar{X}_{tik.}$ = field level harvest loss per acre for harvest loss plot location t

$$\bar{X}_{tik.} = \frac{\sum_{b=1}^{n_b} X_{tikb}}{n_b}$$

where

n_b = number of plots for harvest loss location t,

$X_{tikb} = D_{tikb}E_{tib} + U_{tikb}Z_{tikb}E_{tib}$

where

D_{tikb} = weight of loose and open boll cotton in plot,

E_{tib} = expansion factor,

U_{tikb} = number of large unopen and partially open bolls in plot,

$$Z_{tikb} = \begin{cases} W_{ikb} & \text{if } t = 2 \\ \bar{W}_{ik.} & \text{if } t = 3 \text{ or } 4 \end{cases}$$

Harvest Loss Estimate for Dump Piles

The field level estimator for harvest loss per acre in dump piles is:

$$\bar{X}_{5ik..} = \frac{\sum_{h=1}^{n_i} \bar{X}_{5ikh.}}{n_i}$$

where

$\bar{X}_{5ikh.}$ = harvest loss per acre in dump pile h,

h = dump pile number,

n_i = number of dump piles sampled from field i,

$$\bar{X}_{5ikh.} = (1/A_i)N_iE_iM_{ikh}\frac{\sum_{p=1}^{m_h} X_{5ikhp}}{m_h}$$

A_i = number of acres in field i,

N_i = total number of dump piles for field i,

E_i = expansion factor to pounds per acre,

M_{ikh} = total number of pails in h_{th} sampled dump pile,

m_h = number of pails sampled from the h_{th} sampled dump pile,

X_{5ikhp} = harvest loss per pail in h_{th} dump pile,

= $D_{5ikhp} + U_{5ikhp} \bar{W}_{ik}$.

D_{5ikhp} = weight of loose and open boll seed cotton in p_{th} pail,

U_{5ikhp} = number of large unopen and partially open bolls in p_{th} pail.