

United States
Department of
Agriculture

Rice Objective Yield

1982 Update

Statistical
Reporting
Service

Roberta B. Pense

Statistical
Research
Division

SRS Staff Report
Number AGES 831025

RICE OBJECTIVE YIELD: 1982 UPDATE. By Roberta B. Pense; Statistical Research Division, Statistical Reporting Service, U.S. Department of Agriculture; October 1983. SRS Staff Report No. AGES831025.

ABSTRACT

This study is a follow-up of the 1981 research study in Arkansas. The purposes of this research were to develop objective procedures to estimate rice yield, and to investigate procedures which use multiple regression models to forecast yield early in the season. Based on 1981 and 1982 results in Arkansas, it is possible to estimate yield at harvest. Some potential data collection biases have been identified, but no quantitative measures of bias were made. Several methods of adjusting the estimates for this bias are outlined, although no methods are recommended. It is possible to forecast heads per acre at maturity using early-season counts of stalks or heads. Early-season forecasting equations for weight of grain per head at maturity need more work.

*
* This paper was prepared for limited distribution to the research *
* community outside the U.S. Department of Agriculture. The views *
* expressed herein are not necessarily those of SRS or USDA. *
*

ACKNOWLEDGMENTS

The author wishes to thank Don Von Steen, Stan Hoge, and the office personnel and enumerator staff in the Arkansas State Statistical Office for their comments and suggestions as well as their data collection efforts. Thanks are also extended to the members of Data Collection Branch and Systems Branch who supported this study.

CONTENTS	INTRODUCTION	1
	DATA COLLECTION	1
	ASSUMPTIONS	3
	AT HARVEST ESTIMATION	5
	Yield	5
	Acreage Estimates and Production Estimates	9
	Handling Effect	9
	Plot Location Effect	11
	FORECASTING MODELS	11
	Heads per Acre	11
	Weight per Head	14
	SUMMARY AND RECOMMENDATIONS	16
	REFERENCES	17
	APPENDIX I (Forms)	18
	APPENDIX II (Maturity Code Descriptions)	27

RICE OBJECTIVE YIELD : 1982 UPDATE

Roberta B. Pense

INTRODUCTION

The Yield Assessment Section of the Statistical Reporting Service is developing rice objective yield procedures to be used operationally in 1984 in the five major rice producing states. These states are: Arkansas, California, Louisiana, Mississippi, and Texas. Work began in 1980 with a feasibility study involving nine nonrandomly selected fields in Arkansas (2). In 1981, a sample of 130 randomly selected fields in Arkansas was chosen. This sample provided at-harvest estimates of harvested acreage, yield per acre, and production at the state level. Data were also collected for developing regression models to forecast yield (4). This work continued in Arkansas in 1982 with a sample of 100 randomly selected fields. The objectives in 1982 were basically the same as in the 1981 study. Specifically, the objectives were to:

1. investigate procedures to estimate rice yield at harvest. In 1982 this included looking at the effects on plant growth of repeated visits to the field, and the effects of unit location on yield. It also included examining methods of adjusting yield estimates to account for data collection biases, and evaluating alternative ways to estimate harvest loss,
2. develop regression models to forecast two components of yield -- number of heads per acre, and weight of threshed grain per head at maturity, and
3. elicit comments and suggestions for improvements in data collection procedures from enumerators and state office personnel.

This paper describes changes in data collection procedures and summarizes analysis that was done in 1982. More detail on data collection procedures, previous analysis, and historical background is contained in the SRS staff report "1981 Rice Objective Yield Study"(4).

DATA COLLECTION

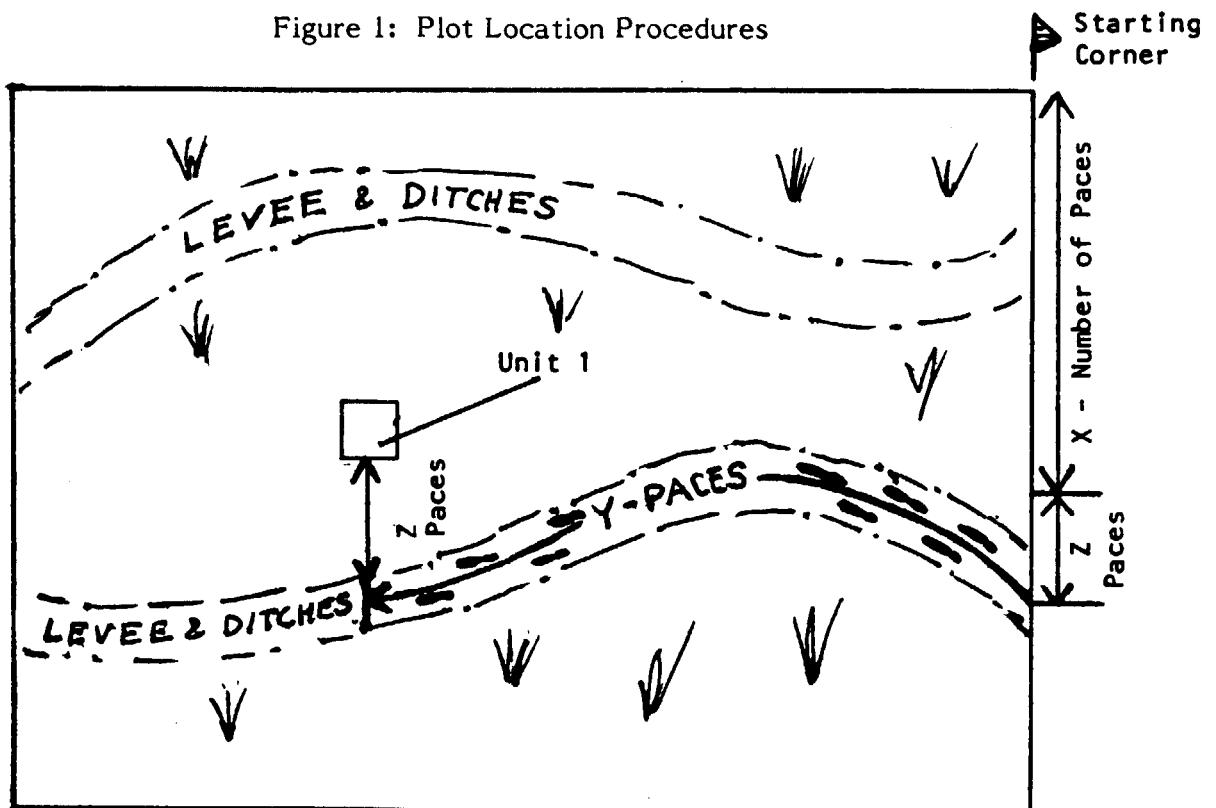
A sample of 100 fields was drawn using the current objective yield sampling scheme (probability proportional to size based on expanded June Enumerative Survey acres planted, or to be planted, to rice). This sample size was adequate to estimate net yield per acre with a coefficient of variation (CV) of less than 5%, based on 1981 variance estimates. Enumerators made field observations on the even-numbered samples at monthly intervals beginning in August as well as immediately before and after harvest. Odd-numbered fields were only visited just before harvest.

Number of plots per sample, plot size, field observations and counts, and clipping instructions were identical to those in the 1981 study. Plots were not located on levees or in ditches in either year. Post-harvest gleaning units were not located in tire tracks. Both of these instructions may introduce bias in yield estimates. However, data collection problems were severe enough that these instructions were necessary.

Procedures for locating Unit 1 changed in 1982, based on suggestions from the enumerators involved in the 1981 study. Previously the enumerators located Unit 1 by pacing x-number of steps along the edge of the field and y-number of steps into the field, where x and y are random numbers based on field size. The damage caused by this method was obvious from the field's edge. The enumerators were concerned about this damage, as were some farm operators. In 1982, the enumerators walked x-number of paces along the edge of the field and visually located the nearest levee ditch. They then counted and recorded the number of paces (z) needed to get to this ditch. They walked y-number of paces down the levee ditch, turned into the field, and walked z-number of paces. Figure 1 illustrates this procedure. While this procedure may not locate the unit in the same place as the 1981 method, the unit location is still random, which is the primary objective of unit location.

Plots tended to be located closer to the edge of the field primarily because the levees curve a great deal. Difficulty with walking in

Figure 1: Plot Location Procedures



muddy ditches is also a factor, but this problem also occurs when walking in a flooded field. Some enumerators even felt that walking down the ditches was easier. Six out of nine enumerators responding to the survey evaluation questionnaire preferred the 1982 unit location procedure because it caused less damage to the field. Those who did not prefer walking down the ditches mentioned curving levees, snakes at levee gates, and confusion caused by the additional instructions as the major problems. It is recommended that the paces into the field in 1983 be 1 1/2 times the paces into the field in 1982 to try to locate plots away from the further edge of the field.

Another change in data collection was that all enumerators were allowed, but not required, to work in pairs. In 1981, some enumerators were assigned to work in pairs and some alone. Many of those assigned to work alone were taking family members or friends with them. The NASDA (National Association of State Departments of Agriculture) cost per sample in 1982 was \$123 as compared to \$142 per sample in 1981. A more detailed cost breakdown between enumerators working alone and those working in pairs is not available. Some decrease in cost per sample was expected in 1982 since enumerators were more experienced. However, it does not appear that allowing enumerators to work in pairs increases costs substantially.

Appendix I contains a copy of all rice objective yield forms. More information on data collection instructions and editing procedures are found in the "1982 Rice Objective Yield Research Study Enumerator's Manual" (7) and the "1982 Rice Objective Yield Supervising and Editing Manual" (8).

ASSUMPTIONS

All of the following analyses assume that there is no difference in yield component estimates between respondents and non-respondents. Table 1 gives some indication of the magnitude of non-response in 1981 and 1982. It is somewhat surprising that the 1982 response rates of 97% and 93% for the initial and post-harvest interviews, respectively, are higher than the 1981 response rates (91% and 81%). Some enumerators had indicated that some respondents in 1981 would refuse in future surveys

TABLE 1: Summary of Responses to Farmer Interview

Response	1981		1982	
	Initial	Post-Harvest	Initial	Post-Harvest
	----- PERCENT -----			
Completed Interview	90.8	80.8	97.0	93.0
No Rice in Tract	3.1	3.1	1.0	1.0
No Rice in Sample but Rice in Tract	1.5	1.5	0.0	0.0
Refusal/Inaccessible	4.6	11.5	2.0	6.0
Missing	0.0	3.1	0.0	0.0
Total	100.0	100.0	100.0	100.0

because of field damage. However, of the 42 samples in 1982 which were also sampled tracts in 1981, only one was a refusal. The tract operator was also a refusal in 1981, so the previous study did not influence his decision.

It is also assumed that little or no bias is introduced into the harvest loss estimate because post-harvest gleaning plots are not located in tire tracks. There should be no bias when the combine used a straw spreader, since harvest loss should be uniformly distributed. If a straw spreader was not used, bias would be introduced. However, in 1982 only 6.5% of the samples were harvested without a straw spreader (7.6% in 1981). Any bias due to these samples should be negligible and constant over the years.

Another assumption is that grain types (long, medium, and short) can be grouped together when building forecasting equations. This assumption is necessary since there are so few observations in the short grain category. Table 2 shows the distribution of samples in 1981 and 1982 by variety. There has been no major shift in variety types so that any effect on regression models due to variety type should be constant over the two years.

TABLE 2: Summary of Varieties by Grain Type

Variety	1981	1982
	-- PERCENT --	
Nortai	1.7	1.0
Total Short Grain	1.7	1.0
Mars	13.6	13.4
Nato	4.2	0.0
Total Medium Grain	17.8	13.4
Labelle	24.6	23.7
Lebonnet	7.6	9.3
Starbonnet	48.3	52.6
Total Long Grain	80.5	85.6
Total All Types	100.0	100.0

All variances were computed using the formula for simple random sampling. This is the procedure used in all operational objective yield programs, but it may not adequately represent the sampling design. The validity of computing the variances in this manner is being investigated and is not addressed in this report. It is assumed that any problems with computing variances in the current method are minimal and consistent with current Agency practice.

AT-HARVEST ESTIMATION

Yield

A final estimate of the net yield per acre at harvest in bushels, adjusted to 12% moisture, was calculated using the data from the final pre-harvest field visit, the laboratory work on the mature samples, and the post-harvest gleanings. The formula for estimating yield per acre is as follows:

Net yield per acre = (Heads per acre x Grain weight per head) - Harvest loss per acre,
where

$$\text{Heads per acre} = \frac{(\text{Number of late boot + emerged} + \text{detached heads in both units}) \times 43560}{(\text{Unit 1 } 5 \text{ row widths} + \text{Unit 2 } 5 \text{ row widths}) \div 5 \times 1.8 \times 3}$$

$$\text{Grain weight per head} = \frac{(\text{Threshed weight of grain in grams}) \times (\text{Threshing loss adjustment factor}) \times (1 - \text{Moisture content})}{(\text{Number of heads threshed}) \times 45 \times 453.6 \times (1 - .12)}$$

$$\text{Harvest loss per acre} = \frac{(\text{Weight of gleaned grain after threshing}) \times 43560 \times (1 - \text{Moisture content})}{((\text{Unit 1 } 5 \text{ row width} + \text{Unit 2 } 5 \text{ row width}) \div 5) \times 1.8 \times 3 \times 453.6 \times 45 \times (1 - .12)}$$

5 adjusts five row widths to one row,
43560 is the number of square feet in an acre,
1.8 is the length in feet of one row (21.6 inches),
3 is the number of rows in one unit,
45 is the number of pounds in a bushel of rice,
453.6 is the number of grams in a pound, and
(1 - .12) adjusts the weight to 12% moisture.

Several methods of estimating harvest loss and net yield were evaluated. These estimates are summarized in Table 3.

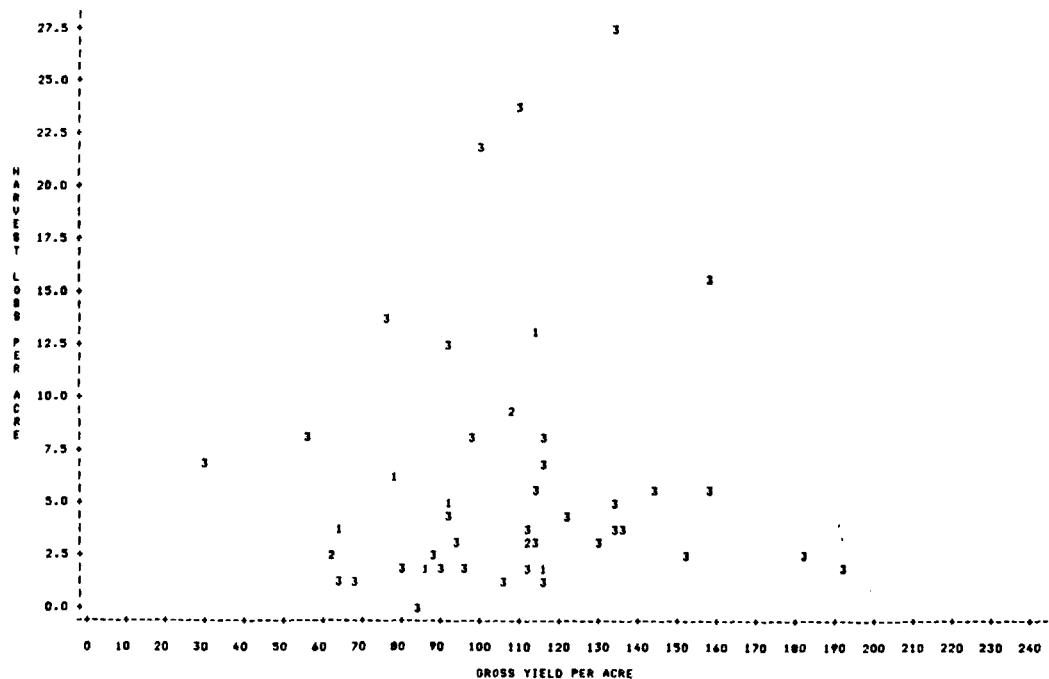
Two methods were used to estimate harvest loss -- an average over all available samples, and an estimate based on stratifying harvest loss by the farmer's reported damage. Ratio and regression estimates of harvest loss were investigated in 1981. These methods did not lead to more precise estimators because there was little correlation between gross yield and harvest loss. This is still true in 1982, as can be seen in Figure 2.

TABLE 3: Summary of Yield Estimates

Variable	n	Mean	Std.Error	CV(%)
Heads per Acre	89	1,227,057	48074	3.9
Wt. per Head (gr.)	89	1.84	0.07	3.9
Gross Yield (bu.)	89	102.8	3.56	3.5
Harvest Loss - Avg (bu.)	47	6.0	0.88	14.7
Harvest Loss - Strat.(bu.)	93	5.9	0.92	15.7
Net Yield - Avg (bu.)	89	96.8	3.67	3.8
Net Yield - Hist. (bu.)	89	95.2	3.72	3.9
Net Yield - Percent Adj. (bu.)	89	93.9	3.93	4.2
Net Yield - Ratio Adj. (bu.)	89	90.2	5.02	5.6
Net Yield - Strat. Adj. (bu.)	89	90.4	3.51	3.9
Farmer Reported Yield (bu.)	93	92.4	1.90	2.1

Figure 2: Plot of Harvest Loss per Acre vs. Gross Yield per Acre (bu/acre)

Symbols are:
 1 = damage affecting harvest loss
 2 = damage during the growing season
 3 = no significant damage



The average estimate of harvest loss is based on the even-numbered samples which received a post-harvest gleaning visit. Since this is a random subsample of the sample, no additional bias is introduced into this estimate. The second estimation procedure involved stratifying harvest loss based on the farmer's reported damage code. Three strata were used: (1) damage, such as lodging, which would affect the harvest loss estimate (2) damage which occurred during the growing season, and

(3) no significant damage. Means were calculated for each stratum based on the available gleaning data. Stratum means were 5.4, 4.3, and 6.3 bushels per acre respectively. The percentage of samples in each stratum was estimated using the available data from the post-harvest interview. The percentages for each stratum were 16%, 14%, and 70% respectively. The standard error for this harvest loss estimator was slightly higher than for the straight average estimate (see Table 3). This indicates a high within stratum variance, which can also be seen in the wide range of harvest loss values for stratum 3 in Figure 2. All subsequent references to harvest loss will be to the straight average estimator.

Net yield per acre was estimated using two methods. They differed only in how harvest loss was estimated for those samples not receiving a post-harvest gleaning visit. The first method expressed net yield as the difference between average gross yield and average harvest loss using only 1982 data (i.e., used the current year's harvest loss estimate for those samples not receiving a gleaning visit). The estimate using this method was 96.8 bushels. A second method used historic harvest loss rather than the current estimate for those samples not receiving a post-harvest gleaning visit. That is, net yield at the sample level was computed for those samples having both gross yield and harvest loss data in 1982. For those samples not receiving gleaning visits, the difference between average gross yield in those samples in 1982 and the average harvest loss in 1981 was used as the net yield estimate. The variance of this "historic" estimator is slightly higher than the first (standard error of 3.72 as opposed to 3.67). For this reason the first "average" method will be used in all subsequent references to net yield in this paper. The "historic" method would probably be used in an operational program, however, since an historic average would be used to forecast harvest loss early in the season.

Since the objective yield plots are not located on levees or in ditches, there is a potential for bias in net yield estimates. Rice grown on levees is more subject to damage from weeds and moisture stress than rice grown within the field. Yield on the levees should be lower than in the rest of the field. In addition, levees were not reseeded for 10% of the samples in 1982. Based on information obtained from the farmer in the initial interview an average of 6.1% of the field is in levees and ditches. This percentage ranged from 1% to 25% of the field, with 65% of the samples in the 1-5% range. Therefore, the objective yield estimate probably overstates yield per acre.

Several procedures can be used to adjust the net yield estimate to eliminate this bias. The simplest procedure is to reduce the estimate by a certain percent. For example, assume the rice yield in ditches is zero and the yield on the levees is not significantly different from the yield in the field. Also assume the acreage in levees is equal to the acreage in ditches. The percent reduction should therefore be the percentage of acreage in ditches. Since an average of 6% of the field is in levees and ditches, let 3% be the percent reduction. This is referred

to as the "percent adjustment" in Table 3. The problems associated with this procedure are that the farmer's levee and ditch acreage estimates may be biased, and the percent adjustment is both arbitrary and subjective in its assumptions.

A second adjustment multiplies the 1982 objective yield estimate by the ratio of the 1981 Crop Reporting Board (CRB) estimate to the 1981 objective yield estimate. This estimate is referred to as the "ratio adjustment" in Table 3. This adjustment must assume that the CRB yield estimate is "true yield". The CRB yield estimate must be treated as a constant (i.e., has no variance) when computing the variance of the ratio adjustment estimate.

A third yield estimate involves estimating yield for levees, ditches, and "within" the field separately. Gross yield per acre would be estimated as $p_1 x_1 + p_2 x_2 + p_3 x_3$, where

- p_1 = percentage of field acreage in levees
- x_1 = yield per acre for those acres in levees
- p_2 = percentage of field acreage that is "within-field"
- x_2 = yield per acre for those acres "within-field"
- p_3 = percentage of field acreage in irrigation ditches
- x_3 = yield per acre for acres in ditches (equals zero and has no variance).

The variance of this estimate would then be approximately $(p_1)^2 \text{ var}(x_1) + (x_1)^2 \text{ var}(p_1) + 2p_1x_1 \text{ cov}(p_1, x_1) + (p_2)^2 \text{ var}(x_2) + (x_2)^2 \text{ var}(p_2) + 2p_2x_2 \text{ cov}(p_2, x_2) + 2 \text{ cov}(p_1x_1, p_2x_2)$.

The harvest loss estimate would be subtracted from gross yield. The variance of gross yield per acre for the levees (x_1) is probably greater than that for the within field (x_2) plots since some farmers reseed levees while other do not. It should also be noted that as the percentage of levees (p_1) increases, the within field yield (x_2) may increase due to improved water management. An additional problem is the accuracy with which p_1 , p_2 , and p_3 are estimated. Acreages are difficult to estimate, especially when they involve small areas. Thus while some data collection biases are eliminated, the variances may increase and another potential bias (acreage estimation) is introduced.

Since no plots were located on levees or ditches, a modification of the third approach was used in 1982. The formula for gross yield was $1/2 (p_1 + p_3) (x_1/x_2) x_2 + p_2 (x_2)$ where p_i and x_i are defined as before. This method assumes that half of the farmer's reported acreage in levees and ditches is in levees. It also assumes that levee yield can be expressed as a percentage of within field yield. Since there is no estimate of this percentage, more assumptions had to be made. It was assumed that (x_1/x_2) could be estimated by solving the following equation for (x_1/x_2) : $1981 \text{ CRB yield} = 1/2 (p_1 + p_3) (x_1/x_2) x_2 + p_2 (x_2) - \text{harvest loss}$, where all the p_i are 1982 estimates, and x_2 and harvest loss are 1981 estimates. This estimate of (x_1/x_2) was then used in the

equation for 1982 to solve for the 1982 CRB yield. Again, this assumes the CRB yield is "true" yield.

None of the net yield estimates in Table 3 are significantly different from each other. The farmer's reported yield was 92.4 bushel per acre, and the CRB yield estimate was 97.5 (adjusted to 12% moisture). The average net yield, unadjusted for bias (96.8), will be used in all subsequent references to net yield since its standard error was the smallest in relation to the mean. It should be noted that this estimate may not have the smallest mean square error. However, since the amount of bias is unknown, the mean square error could not be computed. The average net yield also maintained independence from the Crop Reporting Board estimate, unlike some of the adjusted estimates. Until a validation study can be conducted to eliminate some of the assumptions necessary for adjusting the yield, the simpler "percent" adjustment or a time series chart using average net yield is recommended rather than the more complicated procedures.

Acreage Estimates and Production Estimates

The estimate of planted acres of rice from the June Enumerative Survey (JES) was revised to reflect the acres for harvest. The first revision was done in August, and was based on the ratio of tract acres to be harvested, as reported during the initial interview, to the tract planted acres, as reported on the JES. This ratio was 0.96 in 1982. The revised acreage estimate was 1,359,850 acres, with a standard error of approximately 134,300.

The second revision was based on the field acres harvested as reported on the post-harvest interview. The ratio of this figure to the field acres planned for harvest as reported on the initial interview was 1.007. The revised estimate was 1,370,000 acres with a standard error of approximately 136,400. The final Crop Reporting Board estimate of harvested acres was 1,330,000 acres.

Using the objective yield indications for yield (96.8 bu.) and acreage (1,370,000), the objective yield estimate of total rice production in Arkansas was 132,692,000 bushels. The CRB estimate, adjusted to 12% moisture, was 129,295,000 bushels. The objective yield estimate was therefore 2.6% higher than the CRB estimate.

Handling Effect

Rice in Arkansas is seeded either by using a broadcast method or drilling in 6 inch rows. This, together with flooded conditions early in the season, make it difficult to walk through and make counts in a rice field without some damage to the plants. If the damage is severe enough, the sample plots may not be representative of the "unhandled" areas. In order to investigate the effect of repeated visits, Units 1 and 2 in the even-numbered fields were treated differently. The enumerators located Unit 1 on the first visit and repeatedly observed this unit each month until maturity. They relocated Unit 2 each month.

A Bonferroni paired t-test was used to test the hypothesis that there was no significant difference in counts between the units. This method is described in Timm (6) and the 1981 Rice Objective Yield Study (4). In 1981, no significant differences existed for any month or at maturity. In 1982, a significant difference between the units existed in October and at maturity (see Table 4). Unit 1 contained more emerged heads than Unit 2. This tendency was also present in 1981, which may indicate that damage to the surrounding competition allows more heads to develop fully. The use of two enumerators for all samples in 1982 may have accentuated the effect. However, in 1982, Unit 1 tended to contain more heads (particularly late boot heads) than Unit 2 in August when neither unit had been previously handled. This tendency was not apparent in 1981 and may indicate that the change in unit location procedures affected the plant counts. Since the growing season was approximately the same for both years, it apparently does not explain the year to year difference. Thus, no conclusions can be drawn concerning a handling effect.

Table 4: Summary of Handling Effects ^{1/}

Month	Variable	n	Mean of Unit 1	Mean of Unit 2	\bar{d} (Unit 1-2)	Std. error	t ^{2/}
Aug	Stalks	44	1,427,878	1,396,760	31,118	96,792	0.32
	Late Boot	47	222,386	133,986	88,400	45,613	1.94
	Emerged	47	597,877	641,212	-43,334	82,522	-0.53
	Detached	1	0	0	0	---	---
	Head Wt.	25	0.722	1.004	-0.229	0.0948	-2.42
Sep	Stalks	5	1,228,225	1,147,618	80,607	225,261	0.36
	Late Boot	47	35,777	17,449	18,328	14,660	1.25
	Emerged	47	1,266,465	1,149,006	117,458	67,990	1.73
	Detached	35	1,591	811	780	1,150	0.68
	Head Wt.	47	2.019	2.119	-0.100	0.1051	-0.95
Oct	Stalks	---	---	---	---	---	---
	Late Boot	10	8,067	2,689	5,378	5,378	1.00
	Emerged	10	1,145,749	826,279	319,469	84,457	3.78*
	Detached	9	0	0	0	---	---
	Head Wt.	10	2.473	2.604	-0.130	0.3732	-0.35
Mature	Stalks	--	---	---	---	---	---
	Late Boot	45	0	0	0	---	---
	Emerged	45	1,315,109	1,119,826	195,283	70,019	2.79*
	Detached	45	1,238	631	607	893	0.68
	Head Wt.	46	2.250	2.379	-0.129	0.1237	-1.04

^{1/} Counts are expressed on a per acre basis. Weights are expressed on a per head basis.

^{2/} * indicates the paired means are significantly different at the overall multiple-t significance level of $\alpha=.05$. Hotelling's T^2 tests on appropriate subsets of data yielded same results at $\alpha=.05$ level.

Plot Location Effect

Levees have the effect of subdividing the fields into sub-fields since the water levels and temperatures are relatively constant within levees, and may be different between levees. Rice yields are lower for the sub-fields closest to the water pump because of cold water and water impurities. Yields also tend to be lower in the sub-fields furthest from the pump because of insufficient water.

The enumerators recorded the number of levees from the plot location of Unit 1 to the starting corner. There is no indication of where the water pump is in relation to this corner. While gross yield estimates tended to increase as the plot was located further from the corner, so did harvest loss estimates. The average gross yield for plots in the first four sub-fields was 97.5 bushels per acre as opposed to 111.6 for the other plots. The average harvest loss was 4.3 and 8.1 bushels per acre for the plots in the first four sub-fields and the other plots, respectively. Thus, both gross yield and harvest loss increased for non-corner sub-fields. Net yield is increased, but not as much as gross yield. The study was not designed to examine plot location effect in more detail and no conclusions can be made. Care should be taken to insure plot location does not bias the yield estimates however.

FORECASTING MODELS

Multiple regression models were developed to forecast heads in the sample and weight of grain per head at maturity. Early season head and stalk counts, and early season head weights were obtained for the even-numbered samples in 1981 and 1982. Models should be generated for each maturity category (see Appendix II for a description of maturity categories), but some categories were combined because there were so few observations. The procedures and assumptions for building the models were the same as in 1981. No adjustments were made because of sampling design. Checks for collinearity, influential data points, and heteroscedasticity were made using the regression diagnostics described in Belsley, Kuh, and Welsch (1). Influential data points were deleted when building the models. The "best" model was chosen based on highest R^2 , lowest mean square error, and the least problem with heteroscedasticity or collinearity. See the "1981 Rice Objective Yield Study" report for more details on procedures.

Heads per Acre

The 1981 study showed that the pre-boot and early boot maturity categories could be combined, as well as the milk and soft dough stages. The total number of heads (late boot and emerged) was a better regressor variable than the two head counts individually. These conclusions are still valid based on an inspection of the 1982 data. The independent variables included number of stalks, number of heads, and functions (such as squares, square roots, and logarithms) of these variables. The "best" regression equations are listed in Table 5. Figures 3, 4, and 5 show the regression equations for each of the three maturity category groupings as well as the plots of the data.

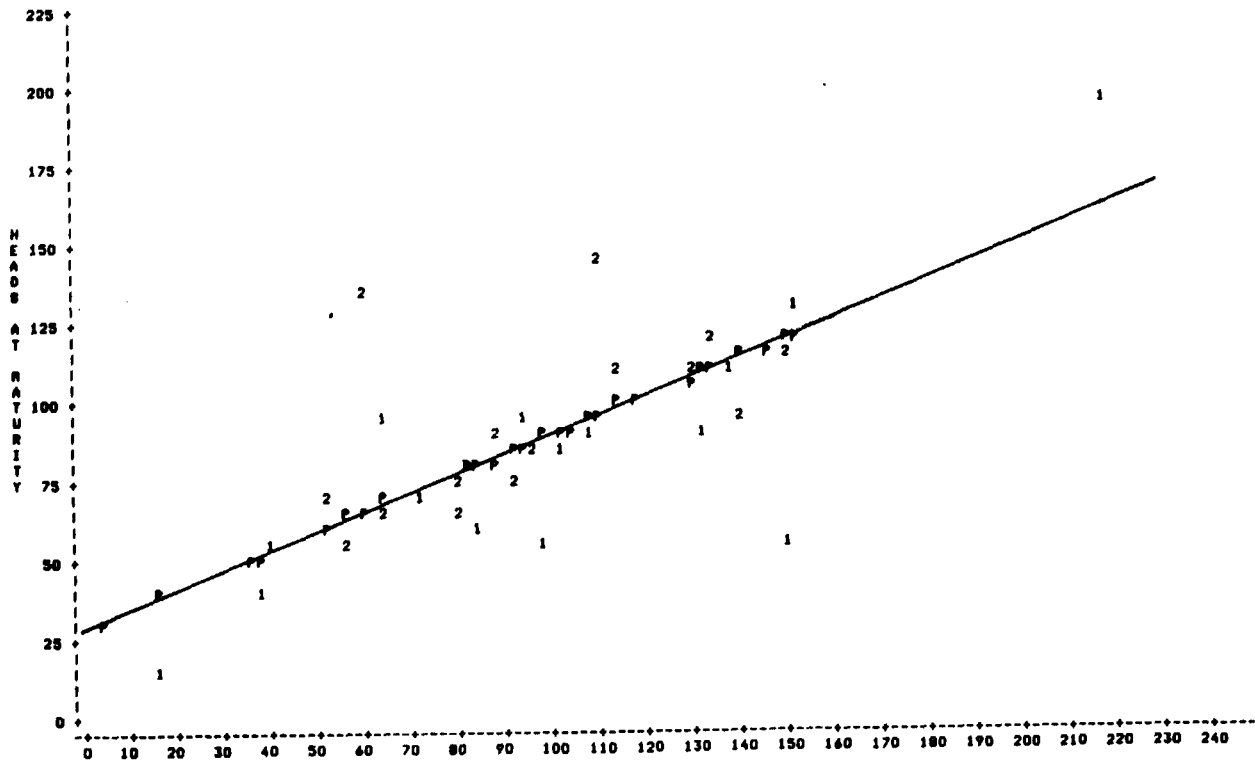
Table 5: "Best" Regression Equations - Number of Heads

Maturity Category	R ²	n	MSE	Equation
Pre-boot & Early boot (1&2)	.53	31*	600.0	29.7606 + 0.6009 (# stalks)
Late boot (3)	.69	49	430.6	35.4037 + 0.7229 (# heads)
Milk and Soft dough (4&5)	.89	33	249.8	-3.7823 + 1.0652 (# heads)

* Some observations were deleted when building the model

A forecasting equation involving number of stalks for the late boot category, had a higher R² (.71) and a lower mean square error (404.9) than the "best" equation for that category. Heteroscedasticity appeared to be more of a problem in the "stalk" equation than in the "head" equation, so the head equation was chosen as best.

Figure 3: Plot of Predicted (P) vs Actual (Symbol is Maturity Category) Heads at Maturity - Maturity Categories 1 and 2



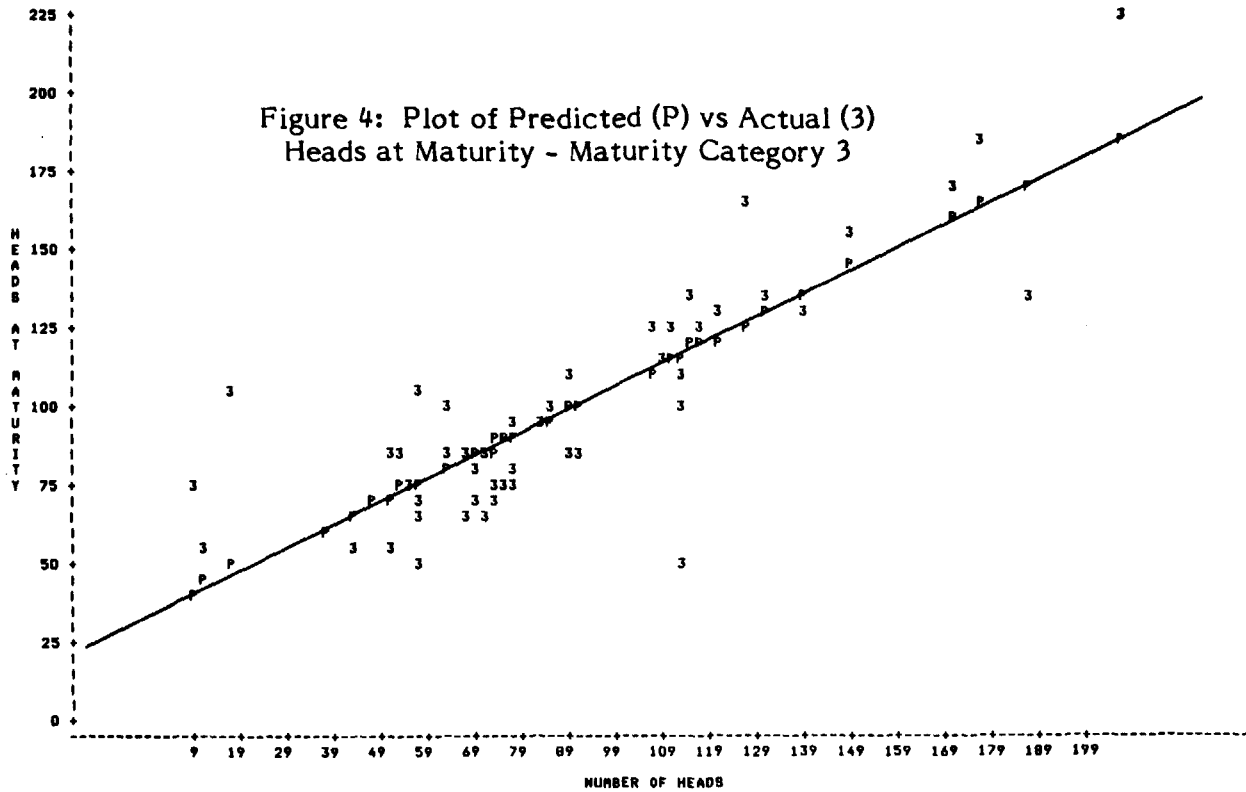
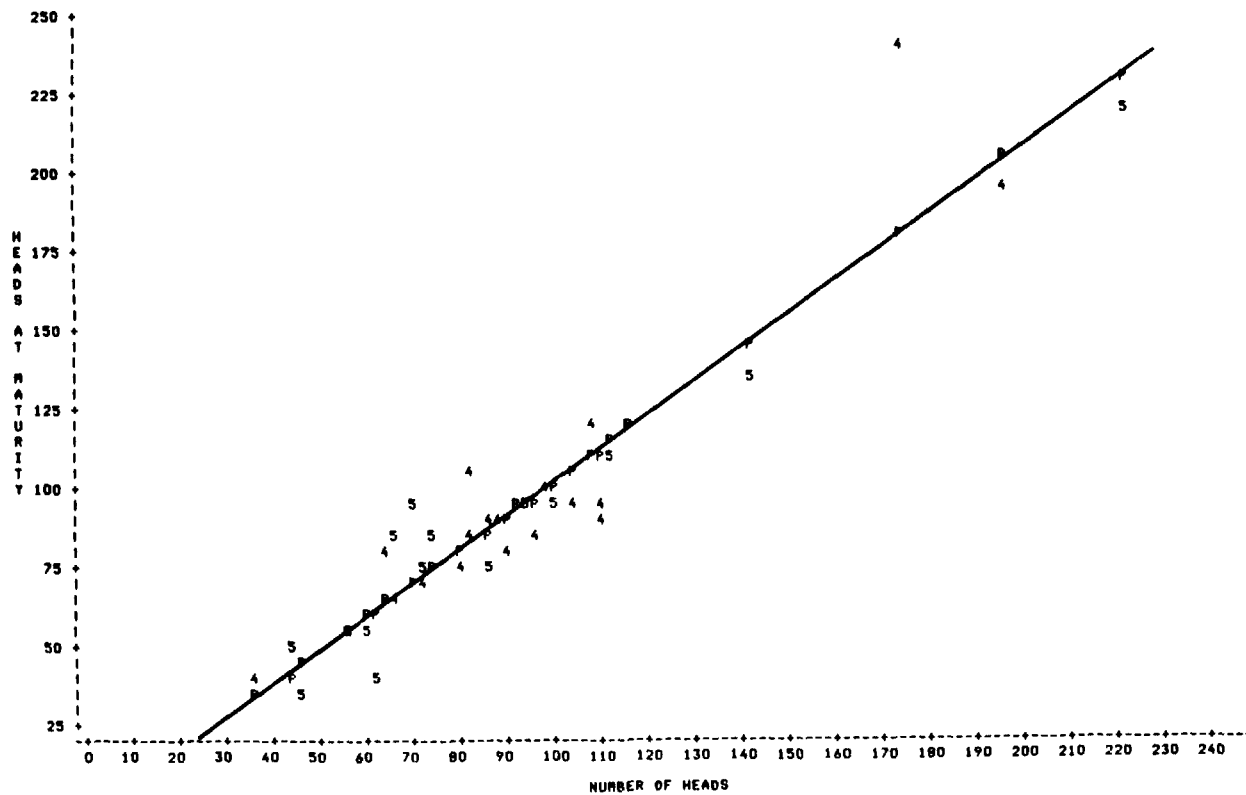


Figure 5: Plot of Predicted (P) vs Actual (Symbol is Maturity Category) Heads at Maturity - Maturity Categories 4 and 5



It should be noted that only data for Unit 1 were used in building the regression equations. This unit was observed throughout the growing season. If a handling effect exists, these regression equations must be adjusted to be applicable to unhandled plots. Alternatively, models could be built using "unhandled" Unit 2 data. Early season and late season relationships probably would not be as strong since different plants are observed each month (10) (11).

Weight per Head

The grain weight per head at maturity in grams, adjusted to 12% moisture, was used as the dependent variable. Early season weights of late boot and emerged heads, and the count of grains per head were used as independent variables. Functions of these variables, such as squares, square roots, and logarithms were also used as independent variables for constructing the "best" regression equation. The weight of heads (late boot and emerged combined) and functions of this variable were also used.

Models were developed by maturity category using both 1981 and 1982 data. However, data for the milk and soft dough stages were combined since there were so few observations and the data plots were similar. The data were also grouped by month rather than maturity category. This grouping was inferior to the maturity category grouping and is not presented.

Table 6 shows the "best" regression equations. Plots of the regression equations as well as the data are found in Figures 6 and 7.

Table 6: "Best" Regression Equations - Weight per Head

Maturity	R ²	n	MSE	Equation
Late boot (3)	.45	44*	0.2116	0.96 + 1.55 (wt/emerged head)
Milk & Soft dough (4&5)	.13	33*	0.3933	1.91 + 0.40 (ln(wt/emerged head))

* Some observations were deleted when building the model.

While the equation for the late boot category is acceptable, the equation for the milk and soft dough category is not. In 1981, the "best" equations for the milk and soft dough category involved the grains per head variable, and had an R² of .46. While heads tended to be lighter in 1982 (2.0 grams as opposed to 2.3 grams) and contained fewer grains (107 as opposed to 126), weight per grain remained constant. This would indicate that grains per head should continue to be a good predictor variable. A plot of the data shows that the additional 1982 data destroys the 1981 relationship. There is no observable pattern or distributional change due to years, however. This is true of all weight per head variables, not just grains per head. This

Figure 6: Plot of Predicted (P) vs Actual (3)
Weight of Grain at Maturity - Maturity Category 3

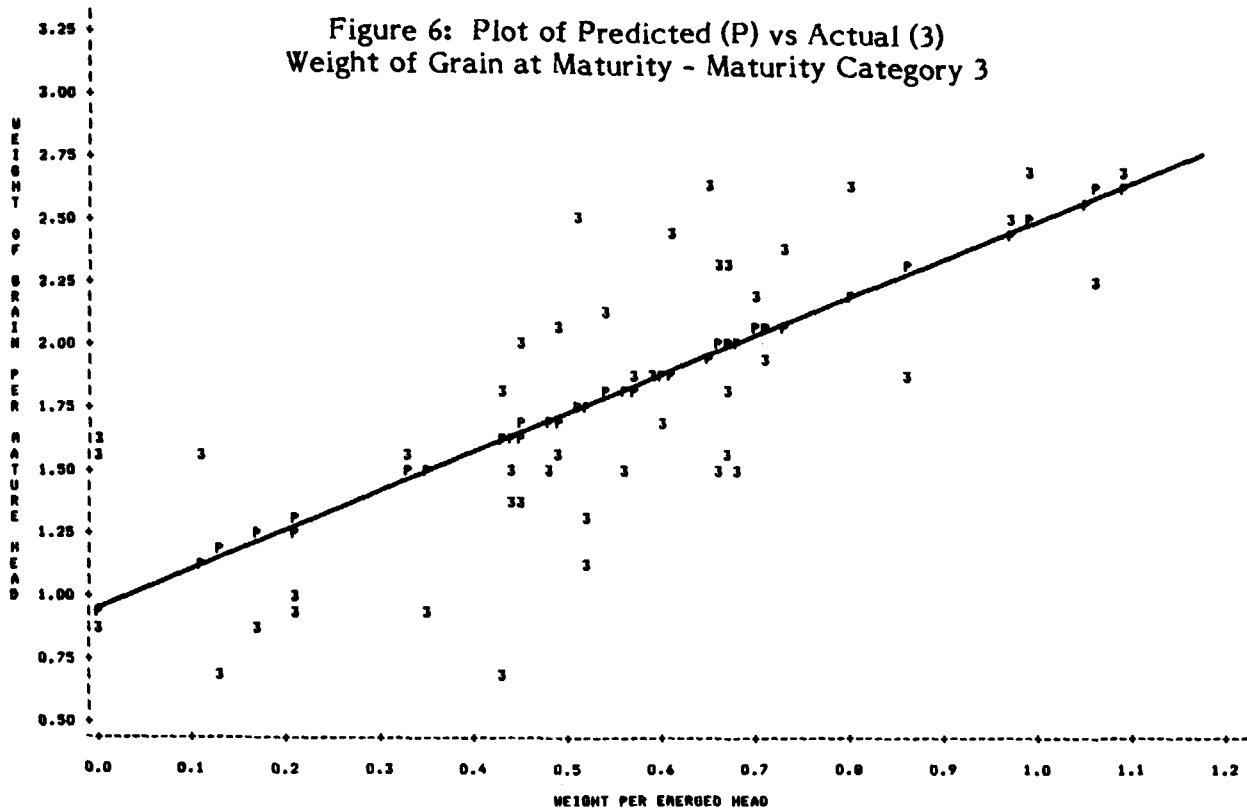
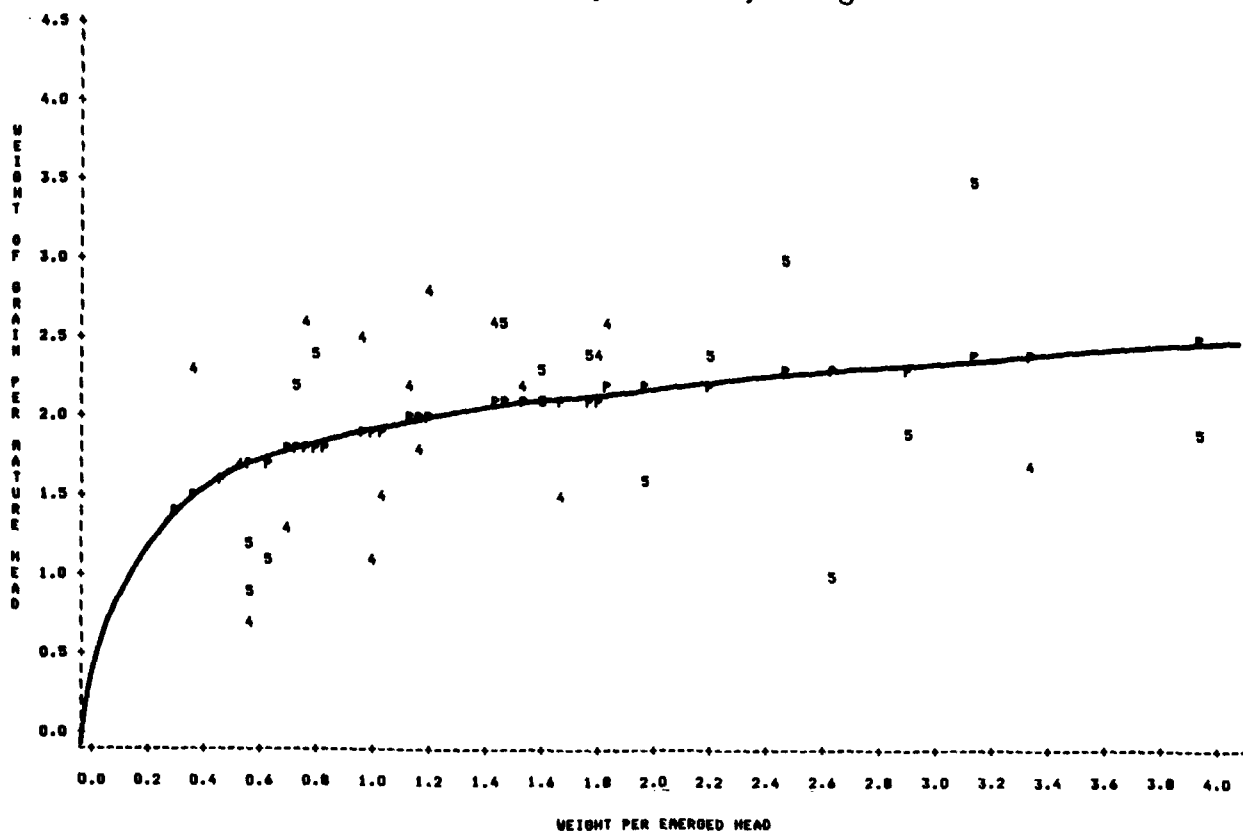


Figure 7: Plot of predicted (P) vs Actual (Symbol is Maturity Category)
Weight of Grain at Maturity - Maturity Categories 4 and 5



fact is particularly important since the SSO used a newer thresher in 1982, and a year effect due to the thresher may have been expected.

SUMMARY AND RECOMMENDATIONS

The assumptions should be summarized before drawing conclusions. It has been assumed that land use stratum and tract or field size have no effect on the estimates when dealing with nonresponse. This same assumption, along with the assumption of no grain type effect, was made when building regression equations to forecast yield components. In addition, variances were computed using the formula for simple random sampling, as is done in operational objective yield programs.

Based on the 1981 and 1982 rice objective yield surveys in Arkansas, the following conclusions can be made:

- 1) It is possible to estimate final yield per acre at harvest using an objective yield procedure. The objective yield estimate was 96.8 bushels per acre in 1982 (CV = 3.8%), which compares favorably with the Crop Reporting Board estimate of 97.5 bushels. In 1981 the objective yield estimate was 110.8 bushels, while the Board estimate was 103.2 bushels.
- 2) Several potential biases have been identified. No plots are located on levees or in ditches, so that yield estimates should be too high. Gleaning plots are not located in tire tracks so that harvest loss estimates should be too high when a straw spreader was not used for harvest. Estimates may be too high or too low depending on the distribution of the sample plots in the sub-fields created by the levees. The 1981 and 1982 studies were not designed to examine these problems and therefore no conclusions or adjustments to the estimates are recommended until these problems are examined in detail. These problems should be addressed in a validation study, where within field relationships are thoroughly examined. A pilot test in Arkansas will be conducted at harvest in 1983 to obtain an estimate of levee yield and its variance and to test data collection procedures.
- 3) No conclusions can be drawn concerning "handling" effect. In 1981, there was no statistically significant effect on yield components due to repeated handling of the plants. In 1982, a significant effect occurred at maturity, even though the growing season was about the same for both years. It is recommended that this study be continued for another year.
- 4) Early season forecasts of heads per sample unit at maturity can be made using early season counts of number of heads. Currently these models assume that there is no handling effect, so that the models are built using data from plots which were observed at least twice. If a handling effect is present, either the forecasts will have to be adjusted for bias, or the models will have to be

built using the "unhandled" data plots. These models will then introduce measurement errors in the independent variables since the observations are not made on the same plants, and the early season and late season relationships are not expected to be as strong.

- 5) Regression equations to forecast weight of grain per head at maturity do not look very promising. Considering the data collection costs, historical averages may be more efficient even though they do not reflect current year situations. It is recommended that data be collected an additional year since the relationships changed so much from 1981 to 1982. Other methods of forecasting grain weight should be investigated.

REFERENCES

1. Belsley, David A., Edwin Kuh and Roy E. Welsch, Regression Diagnostics, John Wiley and Sons, Inc., 1980.
2. Bovard, Gary, "Rice Objective Yield Feasibility Study", memorandum, February 12, 1981.
3. Cochran, William G., Sampling Techniques, John Wiley and Sons, Inc., 1963.
4. Pense Roberta B., 1981 Rice Objective Yield Study, Statistical Research Division, SRS, USDA, Staff Report No. AGES821221.
5. Steel, Robert G.D., and James H. Torrie, Principles and Procedures of Statistics, McGraw-Hill Book Company, Inc., 1960.
6. Timm, Neil H., Multivariate Analysis with Applications in Education and Psychology, Brooks/Cole Publishing Company, 1975.
7. Statistical Reporting Service, 1982 Rice Objective Yield Research Survey Enumerator's Manual, May 1982.
8. Statistical Reporting Service, 1982 Rice Editing Manual, May 1982.
9. Texas Agricultural Experiment Station, Six Decades of Rice Research in Texas, The Texas A&M University System, 1975.
10. Wigton, William H. and Fred B. Warren, Using Objective Measurements of Plant and Soil Characteristics to Forecast Weight of Grain per Head for Winter Wheat, SRS, USDA, November 1971.
11. Wood, Ronald A., Grain Sorghum: A Preliminary Forecast Model, SRS, USDA, September 1972.

Form A: RICE (Cont'd)

Items 3 to 10 apply to the SAMPLE FIELD ONLY.

If no Rice is intended to be harvested for grain in the designated sample field, BUT a NEW field to be harvested for grain is listed in Table A, this new field then becomes the sample field to enter in Item 3 and Item 4.

3. Acres of Rice to be harvested for grain in Sample Field Number _____ Acres
4. What percentage of the _____ acres in Sample Field Number _____ is in levees and ditches? PERCENT
5. What variety of Rice did you seed in this field? ... OFFICE CODE
6. Is this rice, short grain (1) medium grain (2) long grain (3) ENTER CODE
7. Was this field sown by: Broadcast = 1 Drill = 2 ENTER CODE
8. Did you reseed levees? YES = 1 NO = 2 ENTER CODE

9a. Even Numbered Samples

"With your permission I will now go out to the field and mark off two small units to be used in making stalk and head counts."

"I will return to the units each month until harvest to make counts and clip a few heads to determine their weight and size. Would that be all right?" YES NO

b. Odd Numbered Samples

"With your permission I will return shortly before harvest and mark off two small units. I will make counts and clip a few heads to determine their weight and size. Would that be all right? YES NO

10. "After you have finished harvesting this field, I will return to ask you about production. It will be appreciated if you can keep a record of the total amount of rice harvested from this field."

IMPORTANT: Review this form for completeness. Record ending time and sign name. Transfer necessary data from Item 3 to Form D, Item 1.

- Ending Time (Military Time)
- STATUS CODE
-

Enumerator _____

FORM B: RICE YIELD COUNTS - 1982

SURVEY MONTH CODE

YEAR, CROP, FORM, MONTH (1-4)
213

August 1 - 1
 September 1 - 2
 October 1 - 3
 November 1 - 4

UNIT LOCATION

	Paces Levee from No. Levee			
	UNIT 1		UNIT 2	
Number of paces along edge of field				
Number of paces into field				

Date (_____) ... 371

Starting Time (Military Time)

Is this the same unit that was laid out last month?

UNIT 1	UNIT 2
Yes <input type="checkbox"/> No <input type="checkbox"/>	No <input type="checkbox"/>

Check NO if this is the first visit to lay out unit 1 or if this is unit 2.
 Copy the information on "levee number" and "paces from levee" to the sample kit envelope and all other B forms.
 For unit(s) checked: Yes - skip to Item 2.
 No - complete Item 1.

1. Width across 5 row spaces (measure distance from stalks in Row 1 to stalks in Row 6)

UNIT 1	UNIT 2	UNIT 1 #Levees
301	303	300

2. STAGE OF MATURITY: (Circle one stage code for each unit)

Maturity Stage	Pre-Boot	Early Boot	Late Boot or Flower	Milk	Soft Dough	Hard Dough	Ripe
UNIT 1	300 1	300 2	300 3	300 4	300 5	300 6	300 7
UNIT 2	302 1	302 2	302 3	302 4	302 5	302 6	302 7

If the highest maturity code of either unit is Code 1 through Code 4 start counts with 3.

If the highest maturity code of either unit is Code 5, 6 or 7, start counts with 4. For codes 6 or 7, first see Items 7 and 8.

COUNTS WITHIN UNITS

3. Number of stalks (stems) in row

4. No. of heads in LATE BOOT

5. a. Number of emerged heads on all stalks

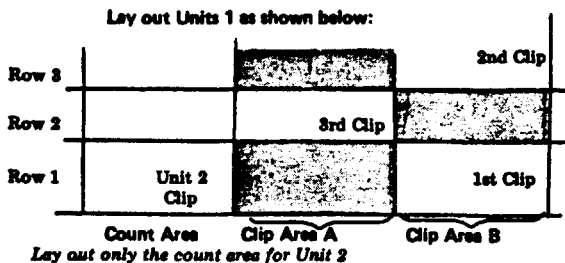
b. No. of detached heads in UNIT (complete ONLY on FINAL PRE-HARVEST VISIT)

UNIT 1			UNIT 2		
Row 1	Row 2	Row 3	Row 1	Row 2	Row 3
311	312	313	314	315	316
351	352	353	354	355	356
331	332	333	334	335	336
	341			344	

6. COMMENTS on condition of field and sample units: _____

(See back - CLIPPING INSTRUCTIONS - Ending Time.)

FORM B: RICE (Cont'd)



CLIPPING ORDER

Unit 1 (Item 8)

- First Clipping – Row 1 in Clip Area B
- Second Clipping – Row 3 in Clip Area B
- Third Clipping – Row 2 in Clip Area A

7. If the HIGHEST MATURITY CODE circled in Item 2 for EITHER Unit is:

- (a) Code 1 or 2: SKIP Items 8 and 9. Enter time and sign name.
- (b) Code 3, 4 or 5: Go to Item 8.
- (c) Code 6 or 7: Go to Item 9.

8. WITHIN CLIP AREAS – Make clippings in the designated ROW within Clip Areas of EACH unit following steps below.

Step 1 – MOW (cut stalk within 2 inches of base) all stalks in specified row until 5 Emerged Heads (if that many) are obtained OR until the row is completely mowed. Begin mowing at end of row farthest from count area and mow in direction of count area. Examine each stalk for emerged head as it is mowed; if present, clip stalk one inch below the head. Place the 5 (or less) emerged heads in 8 # bag. Record count on State (yellow) I.D. tag. Also when mowing, clip and count any heads in late boot and place in 5 # bag.

Step 2 – MOW remaining stalks in row. Examine each stalk and determine which ones are emerged heads and which ones are late boot heads. CLIP the stalk one inch below the head. Place the remaining emerged heads in the 8 # bags and the late boot heads in the 5 # bag.

Step 3 – Record the count of the remaining emerged heads and the late boot heads on the State (yellow) I.D. tag.

Repeat steps 1 thru 3 for Unit 2 using different bags for emerged heads and late boot heads than used in Unit 1.

Prepare two I.D. tags. Label all bags with sample and unit number. Seal and place 8 # and 5 # bags in the 8 # bag.

Verify State (Yellow) I.D. tags and attach to outside of 8 # bags.

Check here after placing 8 # bags in a cloth mailing sack addressed to STATE LAB.
ENTER time and sign name.

9. WITHIN COUNT AREAS – Clip and Count all heads in count area of BOTH units following steps below. Use a separate 8 # bag for each unit.

Step 1 – Clip and Count all Heads in Late Boot in Row 1 – Record in Item 4.

Step 2 – Clip and Count all Emerged Heads in Row 1 – Record in Item 5a and place emerged heads in same bag with late boot heads.

Step 3 – Repeat steps 1 and 2 for ROW 2 and 3. – Record counts.

Step 4 – Pick up and Count all Detached Heads on ground in unit and Record in Item 5b. Place in bag with clipped heads.

Record heads clipped in Items 4 and 5 of Form B and on I.D. Tags. Attach one I.D. Tag to each 8 # bag. Check here () after placing bags in cloth mailing sack addressed to STATE. Enter time and sign name.

ENDING TIME (Military Time)

STATUS CODE

Enumerator _____

CODE

372
380
390

FORM C-1: STATE LABORATORY DETERMINATIONS—
1982 RICE YIELD SURVEY – CLIPPING AREA
SINGLE ROW HEAD SAMPLES

MONTH CODE

Aug. 1 1
 Sept. 1 2
 Oct. 1 3
 Nov. 1 4

YEAR, CROP, FORM, MONTH (1-4)
214 _

Date _____ 470 _____
 (Sample Processed)

1. From Identification Tag

	UNIT 1	UNIT 2	Total
a. All Heads (Emerged and Late Boot) Number			401
b. Stage of Maturity Code			402

2. Laboratory Determinations, Subsample of emerged heads (3# Bag)

	UNIT 1	UNIT 2	
a. Heads in sample (5 or fewer)	403	404	
b. Total weight of heads ... (One decimal)	405	406	
<i>Complete 2c for MATURITY STAGES 4 and 5 ONLY.</i>			
c. Total grains	407	408	

3. Laboratory Determinations on All Remaining Heads

a. Emerged Heads (8 # bag):			
(1) Total number, laboratory count	409	410	
(2) Total weight of heads	411	412	
b. Heads in Late Boot (5 # bag):			
(1) Total number, laboratory count	413	414	
(2) Total weight of late boot heads	415	416	

Lab Technician _____

**FORM C-2: REGIONAL LABORATORY
 DETERMINATIONS - 1982 RICE YIELD SURVEY -
 HARVESTED UNIT HEAD SAMPLES**

MONTH CODE

- Aug 1..... 1
- Sept. 1.... 2
- Oct. 1..... 3
- Nov. 1.... 4
- Dec. 1 or later 5

YEAR, CROP, FORM, MONTH (1-4) <div style="font-size: 24pt; font-weight: bold; margin-top: 5px;">215</div>

570

Date _____
 (Sample Processed)

1. From Identification Tag

	UNIT 1	UNIT 2		
a. All Heads (Emerged, Late Boot and Detached) Number			Total Number	501
b. Stage of Maturity Code			Highest Code	502

2. Laboratory Determinations, all clipped heads from Units 1 and 2

a. Unit 1: (1) Total weight of all heads (One decimal)		503	
(2) Heads in sample		504	
b. Unit 2: (1) Total weight of all heads (One decimal)		505	
(2) Heads in sample		506	
c. Total weight of all heads ... 2a (1) + 2b (1) Grams			

Combine all heads from Units 1 and 2.

3. Threshed grain, all heads from Units 1 and 2

a. Weight immediately after threshing ... (One decimal)		507	
YES <input type="checkbox"/> Go to 3b Is item 3a less than 2c? NO <input type="checkbox"/> STOP - Notify Supervisor.			
b. Weight immediately before moisture test (One decimal)		508	
c. Moisture content 1/ ... (One decimal)		509	
d. Threshing loss adjustment factor (One decimal)		510	

1/ If sample weight is too small for moisture test, sufficient grain of known moisture content will be added to the sample so that a moisture test can be made. The moisture content of the sample can then be derived using the following formula:

$$E = \frac{(A + B) D - (B \times C)}{A}$$

- Where
- A = Weight of small sample (item 3b) _____ Grams
 - B = Weight of additional grain required for moisture test. _____ Grams
 - C = Moisture percent of B _____ Percent
 - D = Moisture percent of A + B combined _____ Percent
 - E = Result -- Moisture percent of small sample (enter in item 3c) _____ Percent

Lab Technician _____

FORM D: RICE YIELD SURVEY - 1982
POST-HARVEST INTERVIEW!

MONTH CODE

Sept. 1..... 2
 Oct. 1..... 3
 Nov. 1..... 4
 Dec. 1 or later 5

YEAR, CROP, FORM, MONTH (1-4)
216

Earlier this year, I (or a representative from our office) contacted you and made some counts on small units in one of your rice fields. I would like to know how your crop turned out in this field.

Date (_____)
 Starting Time

1. Enter from (Form A, Item 3)

Sample Field Number (_____) Acres for Grain (_____)

2. How many acres of rice were (or will be) harvested for grain from this field Acres

If Item 2 is different from Item 1, ask Item 3. If not, skip to Item 4.

DO NOT CHANGE ITEM 1.

3. Earlier in the crop year (Item 1) _____ acres was recorded as being intended for harvest as grain. Can you give me a reason for the difference?

4. How many bushels were harvested from these (Item 2) _____ acres? Total Bushels

If operator indicates yield per acre, multiply by harvested acres to determine total bushels, Show your work.

5. Was production determined from weight tickets Yes = 1 No = 2 Enter Code

6. How many bushels do you still expect to harvest from this field Total Bushels

7. Then the total bushels harvested (or expected) from this field is (Items 4 + 6) Total Bushels (_____)

8. What was the moisture content of the harvested rice

9. On what date was or will harvest be completed in this field? _____ OFFICE USE
 (Month and Day)

10. Was this field harvested with a combine equipped with a straw spreader? Yes = 1 No = 2 Enter Code

11. Was there any significant damage in this field from insects, birds, disease, lodging or other causes? Enter Code

If yes, specify the main source(s) of damage _____ Ending Time

STATUS CODE

CODE

Enumerator _____

FORM E: RICE YIELD SURVEY - 1982

POST-HARVEST GLEANINGS

MONTH CODE

- Aug. 1 1
- Sept. 1 2
- Oct. 1 3
- Nov. 1 4
- Dec. 1 or later 5

YEAR, CROP, FORM, MONTH (1-4)	
217	

Date (_____) 770
 Starting Time (Military Time) 771

The post-harvest field gleanings should be completed as soon after harvest as possible, preferably within three days after harvest. If the sample field has been plowed, diaced or pastured since harvest, select an alternate field for gleaning if one is available in the tract.

UNIT LOCATIONS

	Unit 1	Unit 2
Number of paces along edge of field		
Number of paces into field		
Width across 5 row spaces (measure distance from stalks in Row 1 to stalks in Row 6)..... Feet and Tenths	704	705

GLEANINGS (Place all gleanings from both units in one paper bag.)

1. PICK UP IN BOTH UNITS: a. All unthreshed whole heads
 b. All partly threshed heads
 c. All loose rice grains

CHECK ()	CHECK ()
--------------	--------------

FIELD NOTES: If post-harvest observations cannot be made, give reason here. Indicate if alternate field was selected.

Enumerator _____ Ending Time (Military Time) 772

MAIL gleanings in cloth mailing sack and this Form E in addressed envelope to STATE LABORATORY.

REGIONAL LABORATORY DETERMINATIONS

2. Total weight of heads, kernels and chaff in paper bag (One Decimal)	Grams	701
3. Weight of threshed grain (One Decimal).....	Grams	702
4. Moisture content (One Decimal).....	Percent	703
If samples combined for moisture test, show sample numbers combined: _____		780
DO NOT show combined sample weights in Item 2 or 3.		STATUS CODE
Lab Technician _____	Date Analyzed (_____)	710
	CODE	

5.3 Survey Evaluation Form

Please fill out this questionnaire at the end of the survey period. Your comments will be used in planning future Rice Objective Yield Surveys. Please give a great deal of thought to your answers. If you need more space for your answers, write on the back, or attach another sheet of paper.

1. Were the instructions in the enumerator's manual clear? If not which sections need improvement?

2. Do you have any suggestions as to how to improve the count, unit location, or postharvest gleaning procedures?

3. If you worked on the rice survey last year, do you prefer walking down the ditches as was done this year, or walking into the field from the edge of the field (the way it was done last year)? Why?

4. Are the supplies and equipment you were given adequate? If no, what other supplies do you need?

Are there supplies and equipment that you have now that you do not need?

5. Was farmer refusal a problem?

6. Do you have any major concerns with the rice work (safety, field damage, post-harvest gleanings, unit location, etc.)?

APPENDIX II

Maturity Code Descriptions

CODE 1 - PRE-BOOT

This is a general category in which you will record all units where tillers are only an inch or two high, up to where stalks do not indicate any swelling and DO NOT HAVE the definite flag leaf or other evidence of a partly developed head inside the leaf sheath.

CODE 2 - EARLY BOOT

Stalks are starting to joint and joints can be seen easily. A partly developed head may be detected by noting that the stem has started swelling below the foliage leaf. This swelling may also be felt inside the sheath. Be careful not to damage the partly developed head by squeezing the stem or sheath.

In most cases the presence of heads enclosed in the leaf sheath could be verified by going outside the unit and examining stalks that are similar in appearance to the doubtful ones before classifying the unit in the EARLY BOOT stage. Clip a few stalks, unroll the leaf sheath and see whether or not there is a small, partially developed head encased in the sheath.

CODE 3 - LATE BOOT-FLOWER (HEADS EMERGED) INCLUDES WATERY KERNELS

The head has moved up the stem and swelling has occurred above the base of the top foliage leaf. The sheath will split and the head will partially or wholly emerge. The flower stage occurs soon after the head emerges and small blooms or flowers begin to open at the base of the head and blooming progresses toward the tip. For our purpose, consider the unit to be in the late boot or flower stage from the time swelling can be seen or felt above the top foliage leaf until the head emerges and the watery clear liquid in the kernel has begun to turn milky.

CODE 4 - MILK

Kernels are formed in heads. Kernels of grain are soft, moist and milky. When the grain is squeezed, a milky liquid can be observed. The plant is still generally green.

CODE 5 - SOFT DOUGH The grains can be crushed between the thumb and fingernail; the contents of most of the GRAIN are SOFT with ONLY A FEW GRAINS PER HEAD containing any milky liquid.

CODE 6 - HARD DOUGH

The grain is FIRM and though it may be dented by pressure of the thumbnail, it is NOT EASILY CRUSHED.

CODE 7 - RIPE

Ripe -- straw and leaves may be green or partly green but average moisture in grain is about 20%. Grains at base of head may be in hard

dough stage whereas riper grains in upper portions of the head will be relatively hard. Most of grains will have taken on a mature color but there may be a slightly green color on lower grains. The straw, and to a lesser extent the leaves, may remain fairly green when the grain is considered mature.

CODE 8 - BLANK

This maturity code is used for fields with blank areas where the sample fails. There will be no plants in the sample unit.