Kansas Wheat Non-Sampling Error Analysis

by

Michael E. Craig
USDA/ESCS

Manuel Cárdenas
ASA/USDA Fellow
New Mexico State University

New Techniques Section
Research and Development Branch
Statistical Research Division
Economics, Statistics, and Cooperatives Service
U.S. Department of Agriculture

April 1978

Contents

Acknowledgments

We would like to thank Galen Hart, Bill Wigton, and Jack Nealon for reviewing this paper. Thanks also to Barry Ford for theoretical suggestions. Special thanks to Kathy Wall for her efforts in typing this report.

Summary

Results from analysis of data collected in support of our LANDSAT study in Kansas show digitized acres from aerial photos and reported acres from farmer interviews for wheat are not significantly different when expanded to a three strata total. Individual comparisons at the stratum level do show some significant differences between the reported and "true" acreage. Along with large variation, there seems to be a canceling effect at the strata level leading to the non-significant expanded tests. This canceling is especially visible in the June visit difference test where stratum 11 is significant in the positive direction and both strata 12 and 20 are in the negative direction giving non-significant difference for expanded total. Large variation also seems present at the segment level. The individual enumerator tests showed that some enumerators do a significantly better job than others.

The individual strata comparisons using the paired t-test and the Wilcoxon Signed-Rank came up with approximately the same results by strata. Stratum 11 differences were significant using both types of tests. Comparing the estimates made by the four replications to the actual JES estimate for wheat we see that although the estimates do not differ by more than two JES standard deviations, our closest estimate to the "true" JES acreage (the June Digitized Estimate) is less than the JES by about 1.6 times the JES standard deviation.

Use of paper strata for variance estimation was seen to reduce the variance significantly. Standard deviations increased about 27 percent assuming a simple random sample at the land-use strata only (ignoring paper stratification).

Kansas Wheat Non-Sampling Error Analysis

Introduction

The purpose of this study was to analyze effects of non-sampling errors on June Enumerative Survey in Kansas. The data were originally collected for use with the 1976 Kansas LANDSAT project, and this study was initiated during the ground data editing phase. As the edit originally began, the incidence of errors in ground data when compared to the Color IR photography led to the proposal for a non-sampling error study. The proposal was to use a photo-interpretation of the low level infrared photography at the same time as the LANDSAT project to statistically analyze the non-sampling errors found in the segments. Since the LANDSAT project was aimed at classifying wheat data, the editing procedure categorized fields into "wheat" and "other" cover types. Thus the non-sampling analysis was restricted to the wheat cover type only.

A field acreage determination using the color infrared photography was made with a machine process called digitization. This process related field boundary coordinates to a map base, from which area in acres was calculated. See reference [1] for a further discussion.

Considering the digitized (and photo interpreted) acreage to be a good measure of the "true" acreage, the following objectives were defined (for both June and April visits):

1) test for difference between total expanded acreages for reported versus digitized,
2) test difference by strata and enumerator for wheat at the segment level,
3) report estimates and variances given by digitized acres,

4) report average acres wheat per segment and average wheat field size per strata, and

5) study effect on tests using paper strata (as actually done by June Enumerative Survey) versus assuming simple random sample per strata.

Although the original subsample of the JES included segments from six strata (see Appendix 7), for this analysis it was decided to study only strata 11, 12, and 20 because of small sample sizes in the other three strata. Thus the subsample size was reduced from 174 to 156 for this study (See Table 1). Whenever totals are reported they are for the three strata 11, 12, and 20 and not for all strata.

This study covers the following phases:

a) LANDSAT ground data collection procedures

b) data edit

c) JES estimation

d) design and procedures for testing

e) analysis and estimation results, and

f) summary.

Table 1. Kansas Segment Allocation

Strata	Seg, Size (Sq. Miles)	Total No. Segments	JES Sample	LANDSAT Sample	Paper Strata	JES Reps	LANDSAT Reps
11	1.00	25058	170	68	17	10	4
12	1.00	21732	120	48	12	10	4
20	1.00	21284	100	40	10	10	4

LANDSAT Ground Data Collection Procedures

Original ground data collection in Kansas was directed toward the LANDSAT project. The number of segments to be sampled was decided on in such a manner as to reduce impact on the current JES data collection and to reduce respondent fatigue. Segments were chosen from new JES segments and from segments rotated out of the 1975 survey.

From a total of 435 current Kansas June Enumerative Survey segments, 87 were subsampled for the LANDSAT project. Another 87 were added from segments rotated out of the JES after the 1975 survey. The number 87 consists of two replications each from strata 11, 12, and 20 and 1 rep each from strata 31, 32, and 40. Kansas ESCS enumerators collected data on forms designed by the New Techniques Section with assistance from the Kansas State Statistical Office. Low level color infrared (IR) photography was taken and prepared by the Remote Sensing Institute of the South Dakota State University. These photos were on a scale of 5.25 inches to 1.0 mile. Flights for the segments occurred during the period of May 1 to May 8,1976; between the dates for the two enumeration visits.

Enumerated data were collected on two visits per segment. The first visit was made during the April 12 to May 3, 1976 period; called the April visit. The second visit was made during the May 21 to June 21, 1976 period; called the June visit.

For fields in the selected segments, enumerators collected such items as total field and crop acreage, crop or land use cover, intended use of crop fields, field appearance, and date of harvest. For the current year JES segments, the JES crops section (Part A) was used in June, along with a supplement form to record field appearance codes. For all other segments, a form similar to the JES (Part A) was used which was better suited to recording

appearance and condition codes. (See Appendices 1 and 2 for example forms). The enumerators had only the ASCS photos during visits, not the infrared.

The segment data was then checked in at the Kansas SSO and mailed to the New Techniques Section for editing and keypunching.

Data Edit

Data was edited with both the LANDSAT project and the non-sampling study in mind. It was decided to use digitized acres as truth for LANDSAT and to code reported acres as needed for the non-sampling analysis.

As soon as both ASCS photos and the color IR were received in the Section, the Technical Support staff transferred field and segment boundaries from ASCS to the color IR photography. These boundaries and field numbers were transferred as reported by the enumerators with no attempt made to interpret them. There were 11 segments with unusable or missing IR photos.

Originally, editing was started by hand on enumerator data with respect to all cover types, not just "wheat" and "other". After editing about 1/4 of the data this way, it was concluded that the edit procedure as then described was very unwieldy. At that point, the procedure was changed to edit only wheat data with all other fields called "other". The data already edited was reedited under the new procedure.

Data were edited for reasonableness a segment at a time. All available data were pooled to get the cleanest possible data set for the LANDSAT analysis. At this point, field boundaries were photo-interpreted and compared with the ground data. Acreage and appearance data were then coded on keypunching forms using the instructions as shown in Table 2.

Table 2. Ground Data Edit

Photo Interpreted Cover	Reported Cover	Reported Acres	Digitized Acres
Wheat	Wheat	as rept'd	as digit'd
Wheat	Other	0	as digit'd
Other	Wheat	as rept'd	0
Other	Other	0	0

By using county maps with JES segments drawn in, the segments were located on USGS quad maps wherever possible. Segment and field boundaries were then digitized to a map base using the EDITOR system [1] on the BBN Network. From this system, very precise acreage measurements were available for individual fields. These field measurements were then transferred from BBN to the Washington Computer Center (WCC) and stored on disk.

Coded ground data were inputted into a Statistical Analysis System (SAS) data set and underwent a preliminary machine edit. Digitized acreage for the individual fields was then merged with reported data to make field level records containing April and June reported and digitized acreages, field appearance codes, strata, paper strata, segment number, and visit dates. Some editing for valid appearance codes and ratio of acreages was done in this step. Total digitized segment acreage was compared to total planimetered segment acreage also and any discrepancies were checked. Updates were made on the field level data using the SAS EDITOR Procedure interactively to get a final, "clean" field-level data set.

Wheat fields were then summed to the segment level and merged with a segment master file. The segment level records then contained county number, reported and digitized acreage, appearance code, enumerator I.D. code, enumerator visit dates, total segment acres digitized and planimetered, and ratios of total acres. Using ratios of total acreage, difference (reported - digitized) and scatter plots a final edit and update was then performed to get a "clean" segment-level data set for analysis.

June Enumerative Survey Estimation - Current Procedure

An important source of crop acreage data used by the Economics, Statistics, and Cooperatives Service (ESCS) is the area sample. The completeness of the area frame for sampling leads to theoretically unbiased estimates of population values.

The ESCS area frame consists of strata based on agricultural land use. Each stratum is split into count units [2], and a number of segments is allotted to each land use stratum. Then the number of paper strata is determined, where paper strata may be defined as a group of contiguous count units (or sampling units) thereby creating a geographic stratification. The selection of sample units was done using simple random sampling within paper strata with replication.

Table 3. Kansas JES Expansions (Wheat)

Strata	JES Estimate	JES Std Error	JES CV
11	6276706	235973	3.760
12	4698976	214185	4.558
20	2622947	198003	7.549
ALL	13598629	375185	2.759

The JES currently calculates variances for crops within paper strata or geographic stratification as follows (see Table 3 for actual Kansas JES standard deviations for strata 11, 12, and 20):

$$V_2(\hat{Y}) = \sum_{i=1}^{s} \frac{R_i^2}{r_i} \cdot \frac{\sum_{j=1}^{k_i} \sum_{m=1}^{r_i} (X_{ijm} - \bar{X}_{ij.})^2}{r_i - 1}, \text{ where}$$

X_{ijm} = segment wheat acreage for the mth replication (segment) in the jth paper stratum in the ith land use stratum, where i = 1, 2,..., s; j = 1, 2,..., k_i; m = 1, 2,..., r_i.

$\bar{X}_{ij.}$ = mean per segment in the jth paper stratum in the ith land use stratum

$$= \frac{\sum_{m=1}^{r_i} X_{ijm}}{r_i}$$

$R_i = \frac{N_i}{k_i}$ = number of segments in the population in each paper stratum, within the ith land use stratum.

N_i = number of segments in population in the ith stratum,

r_i = number of replications in sample for the ith stratum, which coincides, by design, with the sample size in each paper stratum in the ith land use stratum,

k_i = number of paper strata in the ith stratum, and

s = number of land use strata.

Note: The variance $V_2(\hat{Y})$ is unbiased. The above notation is a carry-over from the replicated systematic sampling design proposed by Bill Pratt (see reference [2]). Also note that it assumes the number of segments in each paper stratum within land use stratum is constant.

Another objective of this paper was to study the gain in precision due to paper strata versus using a simple random sample per strata. Note that presently sample allocation into paper strata is done proportionally since for a given stratum, R_i and r_i are both constant. Since the total number of segments are greater than 50, and we have proportional allocation in specfic stratum; Cochran (reference [5]) gives a formula for the simple random sample variance which with substitutions in our notation becomes:

$$\hat{V_{ran}(Y)} = \sum_{i=1}^{s} \frac{1}{n_i k_i} \left[\frac{(n_i k_i - k_i + 1)}{n_i k_i} \sum_j S_{ij}^2 + \sum_j \bar{X}_{ij.}^2 - \frac{1}{k_i} \left(\sum_j \bar{X}_{ij.} \right)^2 \right]$$

where,

$n_i = r_i$ k_i = number of segments in the sample in the i<u>th</u> stratum, and

$$S_{ij}^2 = \sum_{m=1}^{r_i} \frac{(X_{ijm} - \bar{X}_{ij.})^2}{r_i - 1}.$$

The direct expansion estimate of the total wheat acreage is calculated as follows:

$$T_{DE} = \sum_{i=1}^{s} \sum_{j=1}^{k_i} \sum_{m=1}^{r_i} \frac{N_i}{n_i} x_{ijm} .$$

Design and Procedures for Testing

The procedures in this section were set up to test for differences in total expanded acreages, reported and digitized; to test for differences by strata at segment level; and to test for differences by enumerator. Other objectives were to study effects of using paper strata for variance calculation versus assuming simple random sample per strata, to report estimates of total acreage, and to report average size of wheat fields.

In order to test the direct expansion estimates of total acreage for digitized versus reported acres of wheat, the following alternatives were proposed:

H_N: T_{DE} (reported) - T_{DE} (digitized) = 0, and

H_A: T_{DE} (reported) - T_{DE} (digitized) $\neq$ 0.

Since the data consists of a natural pairing at the segment level, the total difference (X_{TD}) may be defined as:

$$X_{TD} = \sum_{i=1}^{s} \sum_{j=1}^{k_i} \sum_{m=1}^{r_i} \frac{N_i}{n_i} (E_{ijm} - D_{ijm}), \text{ with}$$

$E_{ijm} = x_{ijm}$ for enumerated or reported acres,

$D_{ijm} = x_{ijm}$ for digitized acres.

So letting $z_{ijm} = E_{ijm} - D_{ijm}$ and $z_{i..} = \bar{E}_{i..} - \bar{D}_{i..}$, and using variance within paper strate, we get:

$$V_2(X_{TD}) = \sum R_i^2 \left(\frac{1}{r_i}\right) \sum_{j=1}^{k_i} \sum_{m=1}^{r_i} \frac{(z_{ijm} - z_{ij.})^2}{r_i - 1}$$

Using Graybill's [3] Theorem 17.1 on the distribution of a linear combination of Chi-Square Variates we are able to get the following approximate t-statistic (see Appendix 4 for Derivation):

$$t_\eta \doteq \frac{X_{TD}}{\sqrt{V_2(X_{TD})}} \qquad \text{where,}$$

$$\eta = \frac{(V_2(X_{TD}))^2}{\sum_{i=1}^{s} \left[\left(\frac{N_i}{n_i}\right)^2 \sum_{j=1}^{m_i} \frac{(z_{ijm} - \bar{z}_{i..})^2}{n_i - 1}\right]^2}{n_i - 1}$$

For tests not involving the direct expansion estimator, the assumption of a simple random sample per land use stratum was used. Along with the test for an expanded total difference, tests were made for individual strata difference on both visits. Strata difference tests were made two ways: paired t-tests and non-parametric Wilcoxon Signed-Rank tests [4]. Signed-Rank tests were also performed for all enumerators who enumerated 5 or more segments (pooling 11, 12, 20 strata).

Using the Signed-Rank test for paired replicates, let (for a given stratum or enumerator i):

Z_{ij} = Acres Reported - Acres Digitized for segment j with j = 1, 2, ..., n_i and take as our model

$Z_{ij} = \Theta_i + e_{ij}$, j = 1, 2,..., n_i (for given i)

where e's are mutually independent random variables, each from a symmetric, continuous population; then we can test hypothesis about Θ_i, the unknown treatment effect.

To test H_N: $\Theta_i = 0$ versus H_A: $\Theta_i \neq 0$ we calculate the Wilcoxon T_i^+ signed rank statistic and reject H_N if:

$$T_i^+ \geq t(\alpha_2, n_i) \text{ or } T_i^+ \leq \frac{n_i(n_i + 1)}{2} - t(\alpha_1, n_i);$$ and accept H_N

otherwise, for specific i and $\alpha_1 + \alpha_2 = \alpha$ level of significance. We also have for large samples (n approaches infinity), the approximation T* which has an asymptotic Normal(0,1) distribution. The statistic T* is defined as:

$$T_i^* = \frac{T_i^+ - [n_i(n_i + 1)/4]}{[n_i(n_i + i)(2n_i + 1)/24]^{\frac{1}{2}}} \quad \text{for given i.}$$

Note: This testing procedure does not take into account any geographic effects which might be present due to the segments not being randomly assigned to the enumerator. Therefore, care must be taken in interpreting the results as a possible enumerator-geographic confounding effect may be present.

An eyeball look at histograms for each stratum (see Appendix 6) and for all strata pooled together showed near symmetric (and near normal) distributions for the differences between reported and digitized acres. Thus at the strata level, both paired t-tests and signed-rank tests were performed and reported on. At the enumerator level, only signed-rank tests were performed due to lack of enough observations per enumerator.

Average wheat acres per segment and per field were calculated by strata. Per segment averages were calculated for reported and digitized values, while the average wheat per field value is from digitized acres only.

Analysis and Estimation Results

Using the four replications available, expanded totals for the three strata are given in Table 4. Comparing these totals to the JES actual direct expansion totals, we notice that although none of the totals falls outside two JES standard deviations, all totals calculated are less than their corresponding JES expansions. This could indicate either a small, non-significant bias in the JES or a bias in the editing procedure for the replications as edited in the New Techniques Section.

Comparing standard errors as calculated within paper strata versus those calculated ignoring paper strata (see Table 5 and 6) we see an increase of 17 to 34 percent when assuming a simple random sample over the whole land use stratum and thus ignoring paper strata. Coefficients of variation (C.V.) allowing for paper strata are given in Table 7. At the three strata level, the June digitized acreage estimate has a C.V. of 4.6 compared to a C.V. of 2.8 for the actual JES sample. Judging from these values, it was felt that the estimates obtained were very precise for using only four-tenths the data available to the JES.

Table 5. Wheat Standard Errors (Using Paper Strata)

Strata	April Reported	April Digitized	June Reported	June Digitized	JES Actual
11	356985	344793	348564	349449	235793
12	338772	361264	336438	348063	214185
20	355450	351597	352665	347552	198003
ALL	607042	610737	599250	603324	375185

Table 4. Wheat Expanded Totals by Strata

Strata	April Reported	April Digitized	June Reported	June Digitized	JES Actual
11	6121448	6020442	6070374	6008945	6276706
12	4633851	4625249	4543256	4553941	4698976
20	2519281	2479905	2402272	2448724	2622947
ALL	13274580	13125596	13015902	13011610	13598629

Table 6. Wheat Standard Errors (Simple Random Sample Per Stratum)

Strata	April Reported	April Digitized	June Reported	June Digitized
11	465477	459593	459811	457420
12	435909	453099	433867	452796
20	441830	428812	414994	417552
ALL	775822	774858	756232	767207

Table 7. Wheat Coefficients of Variation - C.V. (Using Paper Strata)

Strata	April Reported	April Digitized	June Reported	June Digitized	JES Actual
11	5.831	5.727	5.727	5.815	3.760
12	7.311	7.811	7.405	7.643	4.558
20	14.109	14.178	14.680	14.193	7.549
ALL	4.573	4.653	4.604	4.637	2.759

Expanded difference stratum sample means and variances are given in Table 8. Looking at the data before analysis, we see the only large differences show up in stratum 11, while strata 12 and 20 have relatively small differences, both negative and positive. Performing the expanded difference tests on total acres as described in the design section, only the April test using within paper strata variances is near significance (see Table 9).

Table 8. Sampled Data for Expanded Difference Tests

	APRIL		JUNE	
Strata	Differences	s^2(P.S.)	Differences	s^2(P.S.)
11	274.1	8769	166.7	5385
12	19.0	13497	-23.6	8963
20	74.0	4376	-87.3	2527
TOTAL EXPANDED	148984	2.079×10^{10}	4292	1.314×10^{10}

Table 9. Expanded Totals Difference Tests

APRIL			JUNE		
t	n=d.f.	P-Value	t	n=d.f.	P-Value
1.30	7	.24	0.44	23	.68

Considering individual strata, not expanded totals, both normal theory and nonparametric tests were performed on differences per strata. These tests were against the two-sided alternative where a difference in either direction could be significant. The results in Table 10 show a consistent over estimation of wheat acres using reported acreage data for stratum 11 only. Differences in strata 12 and 20 are not significant using either type of test or visit date.

Table 10. Strata Difference Tests

	APRIL						
	Normal Theory			Nonparametric			
Strata	Mean Diff	Paired t_n	P-Value	T^+	N	T*	P-Value
11	4.0	1.55	.13	1376.0	67	1 48	.14
12	0.3	0.09	.93	543.0	47	-0.22	.83
20	1.9	0.59	.56	342.0	36	0.14	.89
				JUNE			
11	2.5	1.22	.23	1433.5	67	1.84	.07
12	-0.5	-0.13	.90	491.5	47	-0.77	.44
20	-2.2	-0.86	.39	247.0	36	-1.35	.18
	0.4	0.23	.82	5774.5	150	0.21	.83

Average wheat acres per segment ranged from 113 to 244 acres considering both reported and digitized acres for the three strata (see Table 11). For stratum 11, wheat acres average about 242 acres; for stratum 12, about 211 acres; and for stratum 20, about 116 acres. For a pooling of the three strata, acres per segment averaged about 200. Average acres per wheat field ranged from 33.8 to 56.0 acres, with a three strata weighted average of 44.4 acres per wheat field.

Table 11. Average Wheat Acres

Strata	Per Segment Reported April	Per Segment Digitized April	Per Segment Reported June	Per Segment Digitized June	Per Field Digitized June
11	244	240	242	240	56.0
12	213	213	209	210	40.4
20	118	117	113	115	33.8
	202	200	199	199	44.4

Difference tests were also done by individual enumerator. For these tests, strata were ignored to get enough data per enumerator. Results are shown in Table 12. Of the thirty-five enumerators, 20 were tested for either April or June or for both visits. Of these tests, five showed a significant difference for reported versus digitized acreage at the level of significance equal to .20. Some other enumerators were very close to the border line on these tests.

Table 12. Difference Tests by Enumerator

	APRIL			JUNE		
Enumerator	N	T^+	P	N	T*	P
1	5	12	.31	2	*	*
2	9	30	.43	7	18	.58
4	8	15	.74	4	*	*
5	5	7	.99	7	16	.81
6	9	32	.30	6	16	.31
8	6	18	.16	6		.84
9	6	15	.44	6	15	.44
10	5	6	.81	0	*	*
11	8	10.5	.38	9	7	.07
13	7	11	.69	7	1	.03
17	6	7	.56	6	20	.06
18	7	15	.94	2	*	*
19	5	9	.81	1	*	*
20	10	40	.23	6	12	.84
24	0	*	*	6	9	.84
27	5	9	.81	6	17	.22
28	7	8	.38	5	3	.31
29	7	24	.11	8	21	.74
30	9	16	.50	9	16	.50
33	5	10	.62	9	29	.50

NOTE: 5 significant at $\alpha = .20$.

*No tests due to small sample size.

References

1. Ozga, Martin; Donovan, Walter E.; and Gleason, Chapman, P.: "An Interactive System for Agricultural Acreage Estimates Using LANDSAT Data," Proceedings of the 1977 Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana.

2. Pratt, William L.: "The Use of Interpenetrating Sampling in Area Frames," Statistical Reporting Service, United States Department of Agriculture, Washington, D.C., May 1974.

3. Graybill, F.A.: "Introduction to Linear Statistical Models Volume 1," McGraw-Hill Book Co., Inc., New York, 1961, pp. 368-369.

4. Hollander, Myles and Wolfe, Douglas A.: "Nonparametric Statistical Methods," John Wiley and Sons, Inc., New York, 1973, pp. 26-35.

5. Cochran, William G.: "Sampling Techniques", Second Edition, John Wiley and Sons, Inc., New York, 1963, pp. 137-141.

Reporting Board
Statistical Reporting Service
U.S. Department of Agriculture
Washington, D.C.

Appendix 1

Form Approved
O.M.B. Number 40-R2766
Approval Expires 4-30-77

1976 KANSAS SATELLITE CROP INFORMATION SUPPLEMENT

NAME OF OPERATOR: _____ (Last) (First) (Middle Initial)

ADDRESS: _____ (Route or Street) (City) (State) (Zip)

NAME OF OPERATION: _____ (Record if different than operator's name above)

ENUMERATOR _____

DATE OF VISIT: Month/Day SEGMENT NO. _____ TRACT LETTER _____ COUNTY _____

(numeric) __ __ / __ __

					Tract Acres 199
	RECORD FIELD NUMBER.....				
	TOTAL ACRES IN FIELD	828	828	828	828
	CROP OR LAND USE *(Specify)*				
	ACRES IRRIGATED AND TO BE IRRIGATED	101	101	101	101
LAND USE COVER	LAND USE COVER WATER COVER *(Lakes, Ponds, Rivers Etc.)*	102	102	102	102
	DENSELY WOODED COVER	103	103	103	103
	GENERAL WASTELAND *(Farmstead, Roads, Ditches, Etc.)*	104	104	104	104
	SUMMER FALLOWED DURING 1976	847	847	847	847
	PERMANENT PASTURE *(Not in crop rotation)*	842	842	842	842
	BARE SOIL OR PREPARED LAND NOT YET PLANTED	105	105	105	105
	STALK FIELD *(Stalks from last years spring planted crops)*	106	106	106	106
	FIELD APPEARANCE CODE *(See Table 1)*	107	107	107	107
CROP TYPE COVER	CROP TYPE COVER WINTER WHEAT Planted	540	540	540	540
	RYE Planted	547	547	547	547
	OATS Planted	533	533	533	533
	BARLEY Planted	535	535	535	535
	CORN *(No Intentions)* Planted	530	530	530	530
	SORGHUM *(No Intentions)* Planted	570	570	570	570
	SOYBEANS *(No Intentions)* Planted	600	600	600	600
	ALFALFA AND ALFALFA MIXTURES Seeded	653	653	653	653
	HAY – OTHER THAN ALFALFA Kind				
	Acres	65_	65_	65_	65_
	OTHER CROPS Name				
	Acres Planted	---	---	---	---
	INTENDED USES OF CROP TYPE OTHER THAN GRAIN Use *(See Table 2)*	150	150	150	150
	Acres	151	151	151	151
	FIELD APPEARANCE CODE *(See Table 3)*	829	829	829	829
	DATE OF HARVEST: If Field Has Been Harvested in 1976	154 MO/DAY __ __ / __ __	154 MO/DAY __ __ / __ __	154 MO/DAY __ __ / __ __	154 MO/DAY __ __ / __ __
	NOTES ON FIELD CONDITION(S) OFFICE USE	155	155	155	155

Field Number Notes

_____ _____

_____ _____

_____ _____

_____ _____

Table 1

FIELD APPEARANCE CODE FOR LAND USE	
10	GREEN COVER (not in planted crop)
20	BARE SOIL (Prepared land not yet planted)
30	DRIED GRASS (brown pasture or fallow)
40	OTHER (Water, F.S. Feed lots, etc.)

Table 2

INTENDED USE OF CROP TYPE	
01	SILAGE
02	CROP TO CUT FOR HAY
03	CROP FOR SEED
04	PASTURED OR GRAZED
05	ABANDONED – left standing
06	ABANDONED – Plowed
07	OTHER – (Soil Imp., etc.)

Table 3

FIELD APPEARANCE CODE FOR CROP TYPE	
50	BARE SOIL (planted but not emerged)
60	GREEN (emerged with green cover, even if partial)
70	MATURE (turning color or ready for harvest) – (not to be used for green hay)
80	HARVESTED CROP (but not worked or prepared, stubble, cut hay etc.)
90	HARVESTED CROP (land worked or plowed)

Form Approved
O. M. B. Number 40-R2766
Approval Expires 4-30-77

Appendix 2

UNITED STATES DEPARTMENT OF AGRICULTURE
Statistical Reporting Service

Segment _____ Page _____ Of _____

State _____ County _____

Enumerator _____

Date of Visit _____

1976 SATELLITE CROP INFORMATION SUPPLEMENT FOR JES

Complete one line for each JES Section A field

FIELD APPEARANCE CODE FOR LAND USE TYPE

10	Green Cover *(not in planted crop)*
20	Bare Soil *(not intended for crop planting)*
30	Dried Grass *(brown pasture or fallow)*
40	Other *(water, F.S., feed lots, etc.)*

FIELD APPEARANCE CODE FOR CROP COVER TYPE

50	Prepared Land-BARE SOIL-*(worked land including planted but not emerged)*
60	Green Cover *(even if partial)*
70	Mature *(turning color or ready for harvest)-(not to be used for green hay)*
80	Harvested Crop *(but not worked or prepared, stubble, cut hay, etc.)*
90	Harvested Crop *(land worked or plowed)*

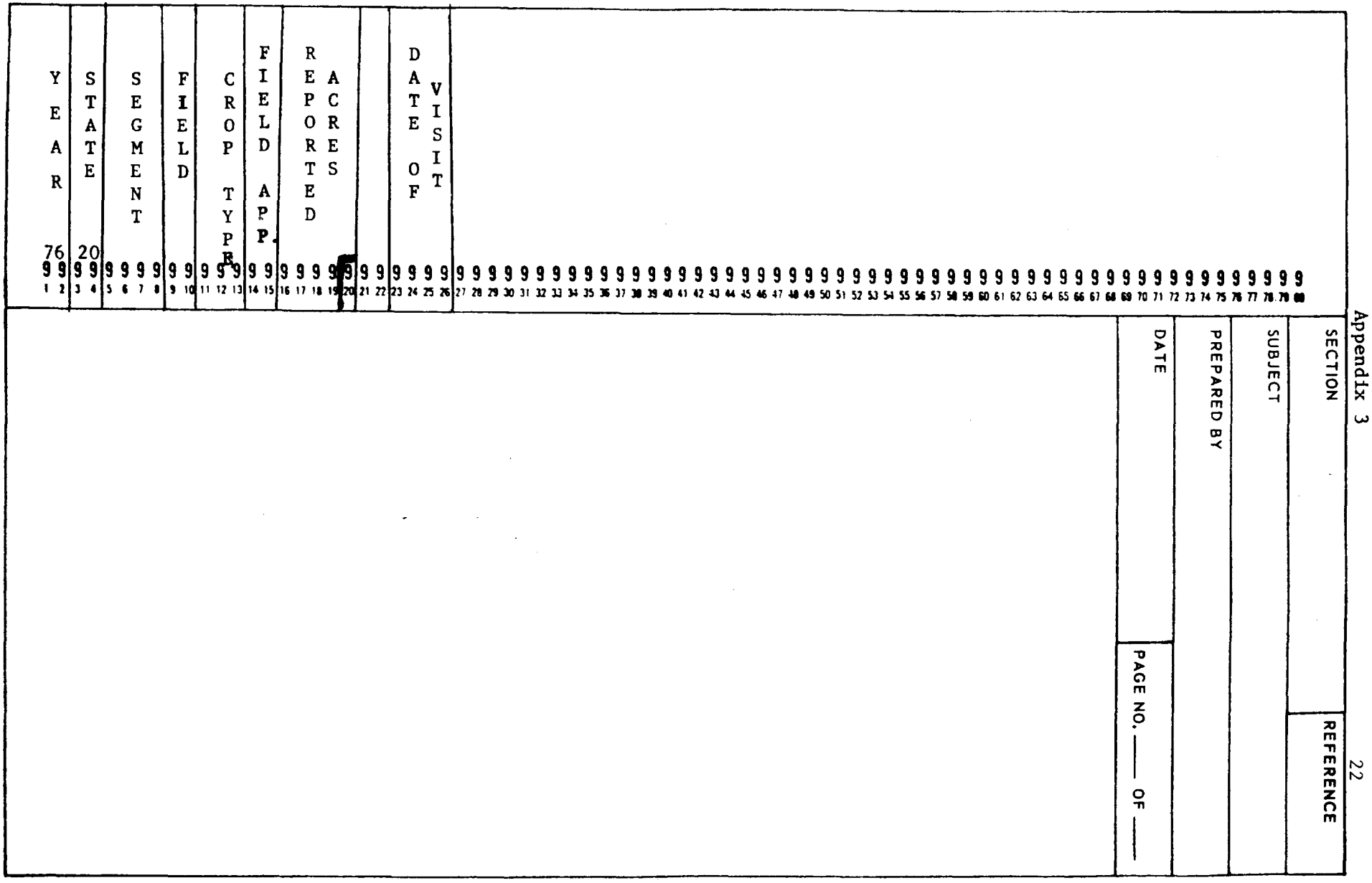
TRACT LETTER	FIELD NUMBER	ACRES IN FIELD	FIELD APPEARANCE CODE	DATE OF HARVEST IF FIELD HAS BEEN HARVESTED	NOTES ON CROP OR FIELD CONDITION

CEF-1011-B
USDA - SRS

System _____

Program _____

RECORD LAYOUT FORM



APPENDIX 4

Derivation of Approximate T-Statistic For Testing
Direct Expansion Totals Difference

Let X_{ijm}, N_i, k_i, X_{TD}, $V_2(\hat{X}_{TD})$, $V_3(\hat{X}_{TD})$, and n: be as defined in the design section. Then the true variance of X_{TD} is a linear function of the variances of the individual strata as follows:

$$V(X_{TD}) = \gamma = g_{11}\delta_{11}^2 + g_{12}\delta_{12}^2 + g_{20}\delta_{20}^2 .$$

An estimate of γ called "g" is of the form

$$g = \sum_i \left(\frac{N_i}{k_i}\right)^2 \frac{1}{r_i} \quad s_i^2 \text{ for } i=11,\ 12,\ 20, \text{ and with}$$

$$s_i^2 = \sum_j \sum_m \frac{(X_{ijm} - \bar{X}_{ij.})^2}{r_i - 1}$$

Thus, we define g_i to be:

$$g_i = \left(\frac{N_i}{k_i}\right)^2 \frac{1}{r_i} \quad \text{for } i=11,\ 12,\ 20.$$

Using Theorem 17.1 [3], then an approximate chi-squared random variable u is given by:

$$u = ng/\gamma = \frac{n(g_{11}s_{11}^2 + g_{12}s_{12}^2 + g_{20}s_{20}^2)}{g_{11}\delta_{11}^2 + g_{12}\delta_{12}^2 + g_{20}\delta_{20}^2} ,$$

and thus an approximate t-statistic is given by:

$$t_n = \frac{X_{TD} - \mu_{TD}}{\sqrt{\frac{V(X_{TD})}{ng/\gamma \cdot \frac{1}{n}}}} = \frac{X_{TD}\sqrt{\gamma}}{g\sqrt{\gamma}} = \frac{X_{TD}}{\sqrt{g}}$$

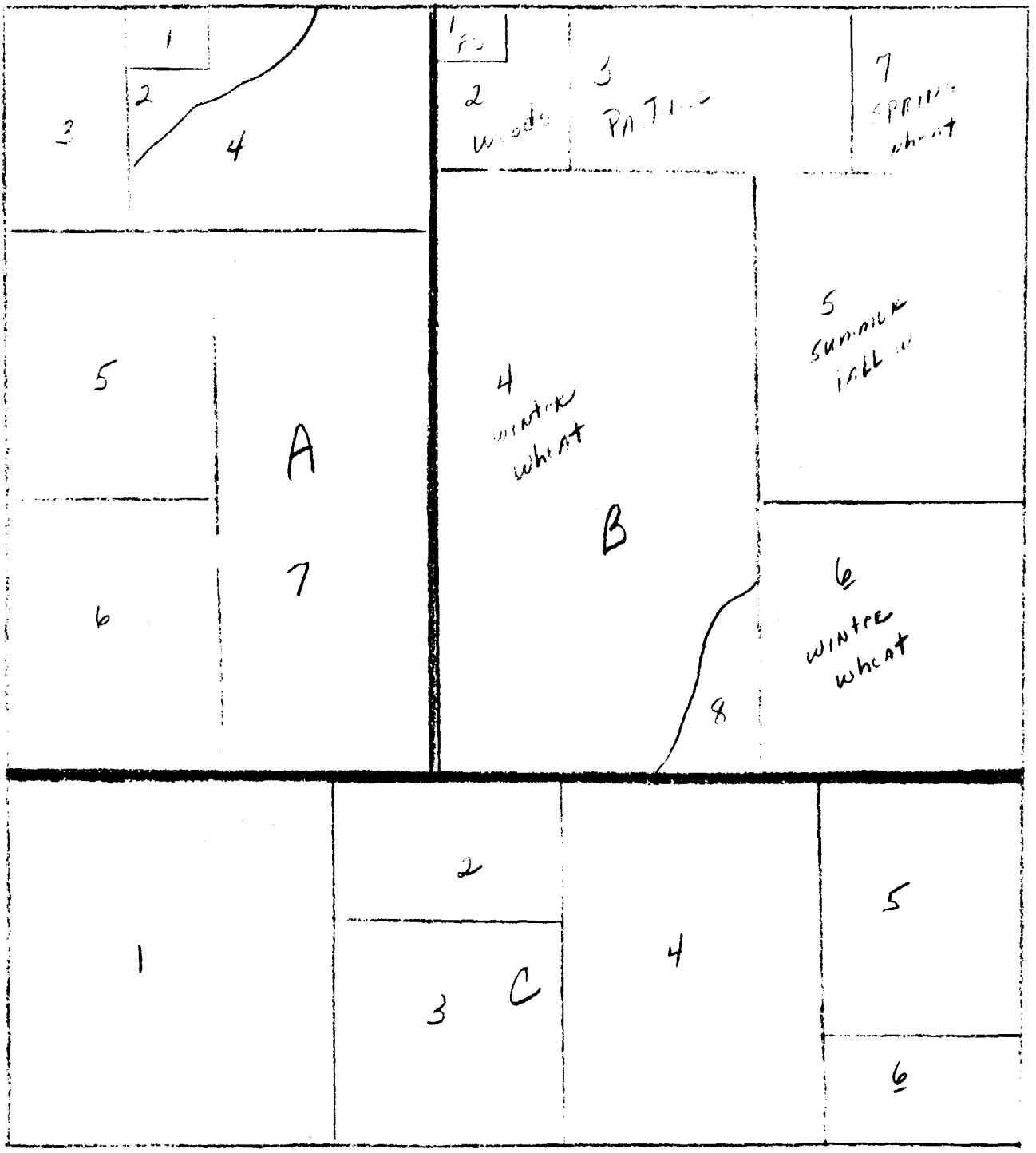
where, $$n = \frac{g^2}{\sum_i g_i^2 \frac{s_i^4}{r_i - 1}}$$

For the variance ignoring paper strata we use the same procedure substituting

$$s_i^2 = \sum_j \frac{(X_{ij} - \bar{X}_{i.})^2}{n_i - 1} \qquad j=1, 2, \ldots, n_i$$

in the corresponding equations using g, g_i, and s_i^2 above.

Appendix 5


5049
1
2
3
4
A
5
6
7
1
2
3
4
B
5
6
7
8
4
winter wheat
5
6
winter wheat
1
2
3
C
4
5
6

Appendix 6

HISTOGRAM OF DIFFERENCES: ALL STRATA

Count 156 Mean 0.358 St. Dev. 19.637

INTERVAL NAME	5 10 15 20 25	FREQUENCY INT.	FREQUENCY CUM.	PERCENTAGE INT.	PERCENTAGE CUM.
-68.000	+X	1	1	0.6	0.6
-64.000	+	0	1	0.0	0.6
-60.000	+	0	1	0.0	0.6
-56.000	+X	1	2	0.6	1.3
-52.000	+	0	2	0.0	1.3
-48.000	+	0	2	0.0	1.3
-44.000	+X	1	3	0.6	1.9
-40.000	+X	1	4	0.6	2.6
-36.000	+XX	2	6	1.3	3.8
-32.000	+	0	6	0.0	3.8
-28.000	+XX	2	8	1.3	5.1
-24.000	+XXXX	4	12	2.6	7.7
-20.000	+X	1	13	0.6	8.3
-16.000	+XX	2	15	1.3	9.6
-12.000	+XXXXXXXXXXXXX	13	28	8.3	17.9
-8.0000	+XXXXXX	6	34	3.8	21.8
-4.0000	+XXXXXXXXXXXXXXXXXXX	20	54	12.8	34.6
0.0000	+XXXXXXXXXXXXXXXXXXXXXXXX	24	78	15.4	50.0
4.0000	+XXXXXXXXXXXXXXXXXXXX	21	99	13.5	63.5
8.0000	+XXXXXXXXXXXXXXXXXX	19	118	12.2	75.6
12.000	+XXXXXXXXXXXXX	14	132	9.0	84.6
16.000	+XXXXXXXXX	9	141	5.8	90.4
20.000	+XXXXX	5	146	3.2	93.6
24.000	+XXXX	4	150	2.6	96.2
28.000	+X	1	151	0.6	96.8
32.000	+X	1	152	0.6	97.4
36.000	+	0	152	0.0	97.4
40.000	+	0	152	0.0	97.4
44.000	+	0	152	0.0	97.4
48.000	+	0	152	0.0	97.4
52.000	+X	1	153	0.6	98.1
56.000	+	0	153	0.0	98.1
60.000	+	0	153	0.0	98.1
64.000	+X	1	154	0.6	98.7
68.000	+	0	154	0.0	98.7
72.000	+	0	154	0.0	98.7
76.000	+	0	154	0.0	98.7
80.000	+	0	154	0.0	98.7
84.000	+	0	154	0.0	98.7
88.000	+X	1	155	0.6	99.4
92.000	+	0	155	0.0	99.4
96.000	+	0	155	0.0	99.4
100.00	+	0	155	0.0	99.4
104.00	+	0	155	0.0	99.4
108.00	+	0	155	0.0	99.4
112.00	+	0	155	0.0	99.4
116.00	+X	1	156	0.6	100.0

Appendix 6

HISTOGRAM OF DIFFERENCES: STRATUM 11

Count 68 Mean 2.451 St. Dev. 16.511

INTERVAL NAME	5 10 15	FREQUENCY INT.	FREQUENCY CUM.	PERCENTAGE INT.	PERCENTAGE CUM.
-45.000	+	0	0	0.0	0.0
-42.000	+	0	0	0.0	0.0
-39.000	+X	1	1	1.5	1.5
-36.000	+X	1	2	1.5	2.9
-33.000	+	0	2	0.0	2.9
-30.000	+	0	2	0.0	2.9
-27.000	+	0	2	0.0	2.9
-24.000	+XXX	3	5	4.4	7.4
-21.000	+	0	5	0.0	7.4
-18.000	+X	1	6	1.5	8.8
-15.000	+X	1	7	1.5	10.3
-12.000	+XX	2	9	2.9	13.2
-9.0000	+XXX	3	12	4.4	17.6
-6.0000	+XX	2	14	2.9	20.6
-3.0000	+XXXXXXX	7	21	10.3	30.9
0.0000	+XXXXXX	6	27	8.8	39.7
3.0000	+XXXXX	5	32	7.4	47.1
6.0000	+XXXXXXXX	8	40	11.8	58.8
9.0000	+XXXXXX	6	46	8.8	67.6
12.000	+XXXXXXXXX	9	55	13.2	80.9
15.000	+XXXX	4	59	5.9	86.8
18.000	+XXXXX	5	64	7.4	94.1
21.000	+X	1	65	1.5	95.6
24.000	+XX	2	67	2.9	98.5
27.000	+	0	67	0.0	98.5
30.000	+	0	67	0.0	98.5
33.000	+	0	67	0.0	98.5
36.000	+	0	67	0.0	98.5
39.000	+	0	67	0.0	98.5
42.000	+	0	67	0.0	98.5
45.000	+	0	67	0.0	98.5
48.000	+	0	67	0.0	98.5
51.000	+	0	67	0.0	98.5
54.000	+	0	67	0.0	98.5
57.000	+	0	67	0.0	98.5
60.000	+	0	67	0.0	98.5
63.000	+	0	67	0.0	98.5
66.000	+	0	67	0.0	98.5
69.000	+	0	67	0.0	98.5
72.000	+	0	67	0.0	98.5
75.000	+	0	67	0.0	98.5
78.000	+	0	67	0.0	98.5
81.000	+	0	67	0.0	98.5
84.000	+	0	67	0.0	98.5
87.000	+X	1	68	1.5	100.0
90.000	+	0	68	0.0	100.0

APPENDIX 7

Kansas Strata Definitions

Strata	Definition
11	80% + cultivated
12	50-80% cultivated
20	15-49% cultivated
31	Urban
32	Urban
33	Urban
40	Range Land (less than 15% cultivated)
50	Non-Agricultural
61	Water